

연구보고서 2016-27

2016년 보건복지통계정보 통합 관리 및 운영



오미애 · 최현수 · 박민규 · 권성훈 · 진재현 · 천미경 · 김솔휘

【책임연구자】

오미애 한국보건사회연구원 부연구위원

【주요 저서】

2015년 보건복지통계정보 통합 관리 및 운영

한국보건사회연구원, 2015(공저)

보건복지통계정보 생산 및 활용 촉진을 위한 마이크로데이터 통합 연계 방안

한국보건사회연구원, 2014(공저)

【공동연구진】

최현수 한국보건사회연구원 연구위원

박민규 고려대학교 통계학과 교수

권성훈 건국대학교 응용통계학과 조교수

진재현 한국보건사회연구원 전문연구원

천미경 한국보건사회연구원 연구원

김솔휘 한국보건사회연구원 연구원

연구보고서 2016-27

2016년 보건복지통계정보 통합 관리 및 운영

발 행 일 2016년 12월 31일

저 자 오 미 애

발 행 인 김 상 호

발 행 처 한국보건사회연구원

주 소 [30147]세종특별자치시 시청대로 370
세종국책연구단지 사회정책동(1~5층)

전 화 대표전화: 044)287-8000

홈페이지 <http://www.kihasa.re.kr>

등 록 1994년 7월 1일 (제8-142호)

인 쇄 처 (주)한디자인코퍼레이션

정 가 6,000원

발간사 <<

보건복지통계는 국민의 삶의 질을 높이기 위하여 수립되는 보건복지 정책의 중요한 기초자료이다. 보건복지에 대한 관심이 높아지면서 보건복지통계 자료 제공에 대한 요구가 증대되고 있고, 시의성 있는 보건복지통계 자료 제공 서비스의 필요성도 점점 증가하고 있다. 2010년 시작된 한국보건사회연구원 보건복지통계정보시스템은 이용자에게 맞춤형으로 통계서비스를 제공하기 위해 지난 2013년까지 4단계에 거쳐 정보시스템 인프라, 기능, 애플리케이션을 보완·개선하였다.

하지만 통계청에서 국가통계포털(KOSIS)를 통해 제공하고 있는 방대한 양의 국가통계 정보와 비교할 때 어떠한 차별성을 지닐 수 있는가에 대한 고민과, 한편으로는 조사데이터의 구축 및 제공에 대한 수요가 점점 더 높아지는 최근의 상황 속에서, 기존의 보건복지통계포털에서 가장 차별화된 영역이었으며 다양한 조사연구를 기반으로 하여 생산된 마이크로데이터에 역량을 집중할 필요성이 대두되었다. 이에 따라 2014년부터 보건복지통계정보시스템을 보건복지데이터포털로 개편하는 작업을 진행하였고, 2015년, 2016년에는 이를 고도화하는 작업을 수행하였다.

2014년 보건복지데이터포털의 주요 개편 내용은 연구 DB, Micro & Macro data, 해외자료(해외 Micro & Macro data, 국제기구통계)로, 데이터에 대한 다양하고 유의미한 정보를 제공하고자 하였으며, 질 높은 데이터 정보 공개 및 데이터 제공 서비스 체계를 구축하였다는 점에서 의의가 있다.

2015년 연구에서는 보건복지통계정보의 통합 제공이 가능한 사이트를 지속적으로 관리·운영하기 위하여 보건복지데이터포털을 이용자 중심으

로 고도화하는 작업을 수행하였다. 화면 디자인을 이용자 중심으로 재설계하였고 기존의 Micro&Macro 부분을 Data Bank로 바꾸어 데이터 저장 및 관리·제공에 대한 역할을 강화하였다. 그리고 데이터 시각화인 인포그래픽을 적극적으로 활용함으로써 통계에 대한 이해도를 높이고, 공간정보통계에서 시도통계 또는 시군구통계를 제공할 수 있도록 하여, 이용자가 직접 지도차트를 생성할 수 있도록 구현하였다.

2016년에는 보건복지통계에 특화된 GIS와 데이터시각화를 활용한 서비스를 확대하여 보건복지통계의 사용자 이해도를 높이하고자 하였다. 그리고 보건복지데이터포털의 이용자들에게 맞춤형 서비스를 제공하고자, 이용자가 원하는 자료를 찾았던 유저들의 패턴을 분석하여 비슷한 자료를 추천해 주는 추천 시스템(Recommendation system)을 구축하였다. 또한, 보건복지데이터포털의 고도화 방향인 마이크로데이터 연계 방안에 대한 연구도 추가하여, 데이터연계 방법론과 데이터연계 평가 및 쟁점을 살펴보았다.

본 연구는 정보통계연구실 데이터연계팀 오미애 부연구위원의 책임하에 최현수 연구위원, 고려대학교 박민규 교수, 건국대학교 권성훈 교수, 진재현 전문연구원, 천미경 연구원, 김솔휘 연구원에 의하여 수행되었다. 바쁘신 중에도 본 보고서를 읽고 조언을 주신 본 원의 황남희 부연구위원, 서울시립대학교 전종준 교수에게 감사의 뜻을 전한다. 또한, 보건복지데이터포털의 고도화 작업을 수행하기 위해 SI업체로 참여한 (주) m114 직원들에게도 감사드린다.

2016년 12월

한국보건사회연구원 원장

김 상 호

목 차

Abstract	1
요 약	3
제1장 서 론	9
제1절 연구배경 및 목적	9
제2절 연구내용	11
제2장 보건복지데이터포털 구축 및 개편 경과	17
제1절 보건복지데이터DB 구축 현황	17
제2절 보건복지데이터포털 운영 현황	25
제3장 보건복지데이터포털 개편 및 신규 콘텐츠 구축	45
제1절 보건복지데이터포털 개편 내용	45
제2절 신규 콘텐츠	72
제4장 보건복지데이터포털 고도화 방향: 추천 시스템	91
제1절 Collaborative recommendation	91
제2절 Content-based recommendation	110
제3절 소결	121
제5장 보건복지데이터포털 고도화 방향: 마이크로데이터 연계	127
제1절 데이터 연계방안에 대한 고찰	127
제2절 데이터 연계 평가 및 쟁점	146

제6장 결론 및 정책 제언	157
----------------------	-----

참고문헌	163
------------	-----

부 록	171
-----------	-----

부록 1. 국내 통계관련 포털 현황	171
---------------------------	-----

부록 2. 보건복지데이터포털 제공 인포그래픽스	204
---------------------------------	-----

표 목차

〈표 2-1〉 보건복지통계정보시스템 연도별 구축 현황	19
〈표 2-2〉 보건복지관련 통계(조사통계 제외) DB수집 누적현황	21
〈표 2-3〉 통계표 중심의 조사통계 DB구축 현황	23
〈표 2-4〉 2014년 보건복지데이터포털 연구 주제별 분류체계	27
〈표 2-5〉 보건복지데이터포털 Data Summary(주제별 I)	31
〈표 2-6〉 보건복지데이터포털 Data Summary(주제별 II)	32
〈표 2-7〉 보건복지데이터포털 Data Summary(조사방법)	33
〈표 2-8〉 보건복지데이터포털 Data Summary(조사범위)	34
〈표 2-9〉 보건복지데이터포털 Data Summary(조사단위)	35
〈표 2-10〉 보건복지데이터포털 Data Summary(조사대상)	36
〈표 2-11〉 보건복지데이터포털 Data Summary(조사주기)	37
〈표 2-12〉 보건복지데이터포털 Data Summary(메타데이터)	38
〈표 2-13〉 보건복지데이터포털 Data Summary(2016년 12월 1일 기준)	38
〈표 2-14〉 보건복지데이터포털 해외 Micro & Macro data 목록	41
〈표 5-2〉 조건부 상관계수 결과	153

그림 목차

[그림 1-1] 2014년 보건복지데이터포털 구축 방향	12
[그림 1-2] 2015년 보건복지데이터포털 개편 방향	13
[그림 1-3] 2016년 보건복지데이터포털 개편 방향	13
[그림 1-4] 보건복지데이터포털 구성도 1	14
[그림 2-1] 보건복지통계정보시스템 목표시스템 구성도	18
[그림 2-2] 애플리케이션 전체 구성도	18
[그림 2-3] 연도별 보건복지관련 통계(조사통계 제외)의 분야별 DB구축 분포도	22
[그림 2-4] 보건복지데이터포털 기본구조 및 주요내용	26
[그림 2-5] 사회보장통계 DB 구축안	29
[그림 6-1] 보건복지데이터포털의 방향성	158
[그림 6-2] 보건복지데이터포털의 구성도 2	159

Abstract <<

Data Portal System Management and Operation 2016 for Health and Welfare Statistical Information

Health and welfare statistics provide an important basis for policies aimed at improving people's quality of life. As public interest in health and welfare increases, the provision of health and welfare statistics to users is becoming increasingly important.

The Korea Institute for Health and Social Affairs(KIHASA) launched the Health and Welfare Statistics Information System in 2010 and since then developed information system infrastructures and applications as part of its effort to provide statistical services tailored to users' needs. Since 2014, the Health and Welfare Statistics Information System has been restructured into what is now known as the Health and Welfare Statistics Portal, which, to set itself apart from Statistics Korea's KOSIS, focuses its capacities on micro-data produced through KIHASA's various surveys. The Portal also features links to relevant international organizations and research bodies.

The Portal was updated to a great extent in 2015, with its main page redesigned in a way to better meet users' needs.

Also, the previous “Micro & Macro Data” was refurbished into “Data Bank” to better archive, manage and provide data. The Portal makes full use of infographics to help improve data visualization. Spatial information is implemented so that the user can create a map chart directly.

This year, we established a recommendation system and a deposit data section in the Data Bank. The recommendation system is useful to predict user preferences. We encourage researchers to consider sharing their own data through the deposit data section. As part of this project, we also looked into ways to data link methods and data link assessment.

1. 연구배경 및 목적

보건복지통계는 보건복지 정책의 중요한 기반이 되는 기초자료로, 국내외적으로 시의성 있고 적절한 보건복지통계정보와 관련 자료를 수요자의 요구에 맞게 제공할 필요성이 점점 커지고 있다. 하지만 보건복지통계정보와 자료 제공에 대한 요구가 높아지고 있는 반면에 정책수립 시 기초자료가 되는 근거중심 통계의 통합 관리 및 서비스 제공이 상대적으로 미흡하다.

그리고 요즘 관심사가 되고 있는 부분이 빅데이터 활용과 마이크로데이터 간의 연계 필요성이다. 보건복지 분야의 조사는 횡단 조사와 종단 조사가 있으나, 패널자료를 제외하고 이질적 표본을 대상으로 하여 조사가 실시되어 조사 간 연계가 어려운 실정이다. 상이한 조사자료 연계는 특정 조사의 한계를 넘어 보건복지 서비스 수요를 파악할 수 있게 하며, 기존 통계의 활용을 통한 미래 예측과 새로운 통계생산의 기반 확보를 가능하게 하기에 데이터 연계 방안도 필요한 연구 분야이다.

보건복지데이터포털에서는 정부 3.0 시대의 국정과제 수행을 위해 필요한 보건복지통계정보 제공 및 역할 모색을 위하여, 2014년부터 보건복지통계정보시스템을 보건복지데이터포털로 개편하는 작업을 진행하였다. 기존의 보건복지통계정보시스템은 통계표를 중심으로 하여 이용자들에게 서비스를 제공하고 있었으나, 통계청에서 제공하고 있는 국가통계포털(KOSIS)의 방대한 국가통계 정보와 차별성을 두고자, 보건복지데이터포털에서는 연구원 고유의 다양한 조사연구를 기반으로 하여 생산된

마이크로데이터에 역량을 집중하였다.

우선, 수요자 중심의 통계정보 및 마이크로데이터 제공 서비스가 가능한 사이트를 구축하기 위하여 한국보건사회연구원에서 수행한 다양한 조사자료에 기반을 둔 마이크로데이터 및 관련 연구 리스트 작성 및 분류체계를 구축하였다. 또한 해외 연구기관 및 국제기구 등과 연계하여 마이크로데이터(패널 등 조사자료)와 매크로데이터(국가별 통계지표 등)에 대한 링크 및 활용방법에 대한 메타정보를 체계적으로 구축하고 제공하여 향후 보건복지통계를 종합적으로 서비스할 수 있는 지속 가능한 포털사이트를 구현하였다. 2015년에는 2014년의 보건복지데이터포털을 개선하는 작업으로, 메인페이지의 구조를 바꾸어 전체메뉴를 한눈에 파악할 수 있도록 하고, 기존의 Micro & Macro 부분을 Data Bank로 바꾸어 데이터 저장 및 관리·제공에 대한 역할을 강화하고자 하였다. 그리고 데이터 시각화인 인포그래픽을 적극적으로 활용함으로써 통계에 대한 이해도를 높였다. 또한, 공간정보통계에서 시도통계 또는 시군구통계를 제공할 수 있도록 하여, 이용자가 지도차트를 생성할 수 있도록 구현하였다.

2016년에는 보건복지통계에 특화된 GIS와 데이터시각화를 활용한 서비스를 확대하여 보건복지통계의 사용자 이해도를 높이하고자 한다. 그리고 보건복지데이터포털의 이용자들에게 맞춤형 서비스를 제공하고자, 이용자가 원하는 자료를 찾았던 유저들의 패턴을 분석하여 비슷한 자료를 추천해 주는 추천 시스템(Recommendation system)을 도입하고자 한다. 마지막으로, 마이크로데이터 연계 방안에 대한 연구로, 데이터연계 방법론과 데이터연계 평가 및 쟁점을 살펴보고자 한다.

보건복지데이터포털은 이러한 고도화작업을 통하여 보건복지 분야 통계자료의 활용도를 향상하고 국민에게 질 높은 통계서비스를 제공할 수 있는 체계를 구축하고자 한다.

2. 주요 연구결과

제2장에서는 1~4차(2010~2013) 기존 포털 시스템, 통계 DB 구축 및 운영 경과를 살펴보고, 2014년, 2015년의 보건복지데이터포털 개편 내용에 대하여 기술하였다. 보건복지데이터포털로의 개편(2014)에서는 보건복지데이터포털의 기본 방향을 제시하였고, 2015년에는 2014년 보건복지데이터포털의 기본 틀을 바탕으로 하여 주요 콘텐츠를 보완하고자 하였으며, 인포그래픽스와 공간정보통계의 신규 콘텐츠를 개발하여 이용자 중심의 서비스를 제공하였다.

제3장에서는 보건복지데이터포털의 올해 신규 내용과 개편된 보건복지데이터포털의 주요 내용을 다루었다. 주요 개편 내용은 각 콘텐츠를 고도화한 부분과 신규 콘텐츠인 추천 시스템(Recommendation system), Deposit data 부분으로 나눌 수 있다.

제4장은 보건복지데이터포털의 고도화 작업으로 추천 시스템의 이론과 장기적으로 확장·적용할 수 있는 방법에 대한 내용을 담고 있다. 추천 시스템은 앞으로도 단계적으로 고도화하여 이용자들에게 향상된 시스템으로 지속적인 서비스를 제공하고자 한다.

제5장에서는 마이크로데이터 통합 연계방안에 대한 전체적인 논평과 관련 논점 및 쟁점을 다루었다. 데이터 연계 방법론, 마이크로데이터 연계 결과에 대한 평가와 통합 마이크로데이터 분석을 위한 가중치, 마이크로데이터 연계의 한계점을 상세히 살펴보았다.

3. 결론 및 시사점

보건복지데이터포털은 연속사업으로, 중·장기적인 계획이 중요한 과제이다. 통계청의 국가통계포털(KOSIS), 사회과학자료원의 사이트와 차별화될 수 있도록, 연구적인 성격을 강화하고 연구자들이 소통할 수 있는 공간을 마련하여 다양한 정보를 제공할 수 있어야 한다. 이를 위해, 보건복지데이터포털의 향후 방향성을 데이터 통합 연계 및 컨설팅으로 제시하고, 서비스 환경을 고도화하며, 지속 가능한 시스템을 관리·운영할 수 있도록 한다. 또한, 운영진은 주기적으로 생산되는 자료의 업데이트와 신규 자료의 등록을 꾸준히 하여 이용자들에게 질 높은 통계서비스를 제공해야 할 것이다. 보건복지데이터포털 콘텐츠와 User Interface는 이용자들의 만족도를 높일 수 있는 중요한 사항이므로 이 부분에 대해 끊임없이 고민하고 노력을 기울여야 할 것이다. 보건복지데이터포털에서 현재 제공하고 있는 서비스 부분과 맞춤형 서비스 지원을 위한 추천시스템의 고도화, 그리고 데이터연계에 관한 다양한 정보를 연결하여 고도화된 시스템을 완성할 수 있다면, 보건복지데이터포털은 보건복지 분야에서 독보적인 존재로 자리매김할 수 있을 것이다.

*주요 용어: 보건복지데이터포털, 추천 시스템(Recommendation system),
Deposit data, 데이터 연계

제 1 장

서론

제1절 연구배경 및 목적

제2절 연구내용

제1절 연구배경 및 목적

1. 연구배경

보건복지통계는 보건복지 정책의 중요한 기반이 되는 기초자료로, 국내 외적으로 시의성 있고 적절한 보건복지통계정보와 관련 자료를 수요자의 요구에 맞게 제공할 필요성이 점점 커지고 있다. 하지만 보건복지통계정보와 자료 제공에 대한 요구가 높아지고 있는 반면에 정책수립 시 기초자료가 되는 근거중심 통계의 통합적 관리 및 서비스 제공이 상대적으로 미흡한 실정이다. 국가 승인통계로 생산된 조사통계의 이용 및 활용률이 점차 증가하는 상황에서 보건복지데이터포털은 보건복지통계정보 및 데이터의 통합 관리로 수요자의 요구에 맞는 정보를 제공할 필요성이 있다.

또한, 보건복지 분야의 조사 자료에 대한 사회적 수요가 증가하고 있다. 마이크로데이터의 이용자가 증가하고 있고, 학문적 활용도 제고를 위한 정보공개 요구가 꾸준히 증가하고 있으며, 이를 위해 통계 승인을 받은 조사의 마이크로데이터를 수집하여 축적하고, 이를 체계적으로 관리, 제공할 수 있는 포털을 구축하고 운영할 필요가 있다. 포털 운영은 대규모 조사사업을 통하여 생산된 원시자료의 활용도를 높여 통계 소비자(전문연구자, 정책담당자 등)의 만족도를 제고할 수 있다.

그리고 요즘 관심사가 되고 있는 부분이 빅데이터 활용과 마이크로데이터 간의 연계 필요성이다. 보건복지 분야의 조사는 횡단 조사와 종단

조사가 있으나, 패널자료를 제외하고 이질적 표본을 대상으로 하여 조사가 실시되어 조사 간 연계가 어렵다. 상이한 조사자료 연계는 특정 조사의 한계를 넘어 보건복지 서비스 수요에 대한 파악이 가능하며, 기존 통계의 활용을 통한 미래 예측과 새로운 통계생산의 기반 확보가 가능하기에 데이터 연계 연구 방안에 대한 연구도 필요하다.

마지막으로, 포털에서는 정부 3.0 시대의 국정과제 수행을 위해 필요한 보건복지통계정보 제공 및 역할을 모색하여야 한다. 보건복지데이터 포털은 포털 운영 관리에 그치는 것이 아니라, 근거기반 연구의 효율화를 위해 데이터 연계와 같은 심도 깊은 연구 및 학술 정보를 제공할 필요성이 있다.

2. 연구목적

기존의 보건복지통계정보시스템은 통계표를 중심으로 하여 이용자들에게 서비스를 제공하고 있었으나, 통계청에서 제공하고 있는 국가통계포털(KOSIS)의 방대한 국가통계 정보와 차별성을 두고자, 연구원 고유의 다양한 조사연구를 기반으로 하여 생산된 마이크로데이터에 역량을 집중하고자 하였다. 이에 따라, 2014년부터 보건복지통계정보시스템을 보건복지데이터포털로 개편하는 작업을 진행하였다.

수요자 중심의 통계정보 및 마이크로데이터 제공 서비스가 가능한 사이트를 구축하기 위하여 한국보건사회연구원에서 수행한 다양한 조사자료에 기반을 둔 마이크로데이터 및 관련 연구 리스트 작성 및 분류체계를 구축하였다. 또한 해외 연구기관 및 국제기구 등과 연계하여 마이크로데이터(패널 등 조사자료)와 매크로데이터(국가별 통계지표 등)에 대한 링크 및 활용방법에 대한 메타정보를 체계적으로 구축하고 제공하여

향후 보건복지통계를 종합적으로 서비스할 수 있는 지속 가능한 포털사이트를 구현하였다. 2015년에는 2014년의 보건복지데이터포털을 개선하는 작업으로, 메인페이지의 구조를 바꾸어 전체메뉴를 한눈에 파악할 수 있도록 하고, 기존의 Micro&Macro 부분을 Data Bank로 바꾸어 데이터 저장 및 관리·제공에 대한 역할을 강화하고자 하였다. 그리고 데이터 시각화인 인포그래픽을 적극적으로 활용함으로써 통계에 대한 이해도를 높였다. 또한, 공간정보통계에서 시도통계 또는 시군구통계를 제공할 수 있도록 하여, 이용자가 지도차트를 생성할 수 있도록 구현하였다.

2016년에는 보건복지통계에 특화된 GIS와 데이터시각화를 활용한 서비스를 확대하여 보건복지통계의 사용자 이해도를 높이하고자 한다. 그리고 보건복지데이터포털의 이용자들에게 맞춤형 서비스를 제공하고자, 이용자가 원하는 자료를 찾았던 유저들의 패턴을 분석하여 비슷한 자료를 추천해 주는 추천 시스템(Recommendation system)을 도입하고자 한다. 마지막으로, 마이크로데이터 연계 방안에 대한 연구로, 데이터연계 방법론과 데이터연계 평가 및 쟁점을 살펴보고자 한다.

보건복지데이터포털은 이러한 고도화작업을 통하여 보건복지분야 통계자료의 활용도를 향상하고 국민에게 질 높은 통계서비스를 제공할 수 있는 체계를 구축하고자 한다.

제2절 연구내용

본 과제는 2010년부터 진행된 연속사업으로 보건복지통계정보시스템과 보건복지데이터포털을 통해 이용자 중심의 서비스를 제공하고자 노력하였다. 그동안의 보건복지데이터포털 구축 및 경과를 살펴보고, 올해의 개편 내용을 정리하여 기술하였다. 또한, 보건복지데이터포털의 고도화

방향인 추천 시스템과 마이크로데이터 연계 방안에 대한 연구를 통해 보건복지데이터포털이 앞으로 나아가야 할 방향성을 제시하고자 하였다.

기존의 보건복지데이터포털 시스템과 서비스 부분은 2장, 보건복지데이터포털의 업데이트 부분은 3장, 보건복지데이터포털의 고도화 방향은 4장, 5장에서 다루었다.

제2장 “보건복지데이터포털 구축 및 개편 경과”에서는 그동안의 보건복지통계포털 및 보건복지데이터포털의 구축 및 운영 경과에 대해 살펴 보았다. 보건복지통계정보시스템은 2010년부터 2013년까지 수행되었으며, 매년 점차적으로 정보시스템의 기능 및 인프라, 애플리케이션 등을 보완하고 개선하는 작업을 진행하였다. 정보제공 범위는 통계표 중심의 통계DB를 구축하여 보고통계 및 조사통계, 마이크로데이터, 통계관련 정보를 제공하였다. 보건복지데이터포털로의 개편(2014)에서는 보건복지데이터포털의 기본 방향을 제시하였고 2015년에는 이를 개선하는 작업을 진행하였다.

[그림 1-1] 2014년 보건복지데이터포털 구축 방향



[그림 1-2] 2015년 보건복지데이터포털 개편 방향



제3장 “보건복지데이터포털 개편 및 신규 콘텐츠 구축”에서는 보건복지데이터포털의 올해 신규 내용과 개편된 보건복지데이터포털의 주요 내용을 다루었다. 주요 개편 내용은 각 콘텐츠를 고도화한 부분과 신규 콘텐츠인 추천 시스템(Recommendation system), Deposit data 부분으로 나눌 수 있다.

[그림 1-3] 2016년 보건복지데이터포털 개편 방향

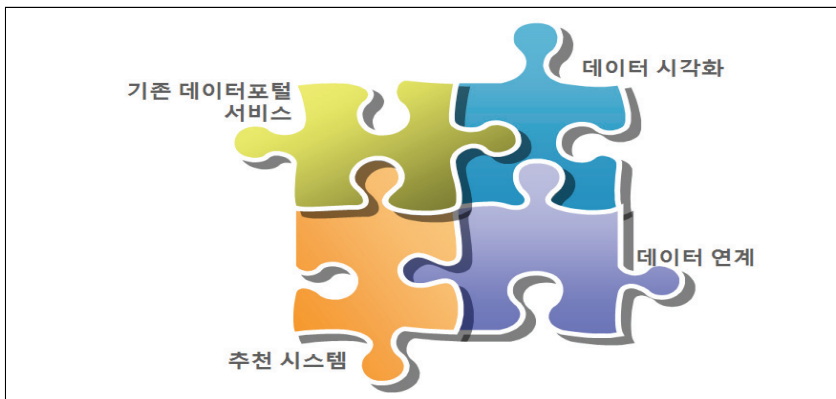


제4장 “보건복지데이터포털 고도화 방향: 추천 시스템(Recommendation system)”에서는 보건복지데이터포털의 고도화 작업으로 추천 시스템의 이론과 장기적으로 확장·적용할 수 있는 방법에 대한 내용을 담고 있다. 제5장 “보건복지데이터포털 고도화 방향: 마이크로데이터 연계”에서는 마이크로데이터 통합 연계방안에 대한 전체적인 리뷰와 관련 논점 및 쟁점을 다루었다.

각 장들은 유기적으로 연관성이 있는데, 기존의 데이터포털 서비스는 축적된 정보를 통해 추천 시스템 구축에 활용되며, 추천시스템으로 인기가 있거나 관련성이 높은 데이터와 연구과제 목록이 도출되면 데이터 연계의 후보군이 되어 연구자들에게 통합 데이터로 제공할 수 있게 된다. 그리고 의미 있는 결과들은 시각화를 통하여 이용자들에게 더 쉽게 전달될 수 있도록 한다.

마지막으로, 제6장 “결론 및 정책 제언”에서는 보건복지데이터포털의 향후 목표를 데이터 통합 연계 및 컨설팅으로 설정하여 나아가고자 하는 방향성을 제시하고, 이에 맞는 서비스 환경을 고도화하여, 지속 가능한 시스템을 관리·운영할 수 있도록 한다.

[그림 1-4] 보건복지데이터포털 구성도 1



제 2 장

보건복지데이터포털 구축 및 개편 경과

제1절 보건복지데이터DB 구축 현황

제2절 보건복지데이터포털 운영 현황

2

보건복지데이터포털 구축 및 개편 경과¹⁾

제1절 보건복지데이터DB 구축 현황

1. 정보화전략계획 수립

경제 성장 및 삶의 질 향상 등 국내 외적 환경의 변화로 보건복지통계 자료 제공에 대한 요구가 증대하면서 통계자료의 효율적 제공 및 관리가 필요함을 인식하게 되었다. 그뿐만 아니라 국가 승인통계 생산의 기반이 되는 가치 있는 조사 자료를 이용하려는 수요가 증가함에 따라 조사 자료의 체계적 관리가 더욱 절실하게 되었다. 이에 보건복지통계정보시스템의 구축 필요성이 대두되었고, 그에 따라 보건복지통계DB의 성공적 구축을 위하여 우선적으로 시스템 구축의 중·장기적인 정보화전략계획(ISP; Information Strategy Planning)을 2009년에 수립하였다.

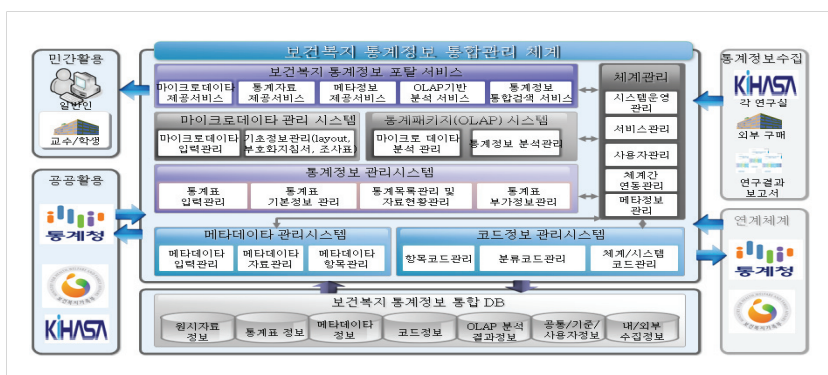
정보화전략계획 수립의 추진목표는 보건복지통계 생산과 관련된 업무 프로세스 및 정보의 흐름에 맞춘 정보화 계획을 수립하고, 정보자원에 대한 표준을 정립하여, 보건복지통계 이용자에게 맞춤형 서비스가 가능한 시스템 구축방안을 제시하는 것이었다. 이에 따라 정보화전략계획 수립을 통해 [그림 2-1]의 목표시스템 구성도가 도출되었다.

보건복지통계정보 통합 관리체계의 목표 구조는 전체 6개 체계와 하부에 21개 서비스시스템으로 구성되었다. 구체적으로 살펴보면, 통계정보

1) 보건복지데이터포털은 연속사업으로 전체적인 구성을 위해 [오미애 외(2015). 2015년 보건복지통계정보 통합 관리 및 운영, 한국보건사회연구원]의 내용을 포괄적으로 인용함.

포털서비스, 마이크로데이터 관리, 통계정보 관리, 메타정보 관리, 코드 관리, OLAP(OnLine Analysis Processing) 등 6개 체계를 중심으로 하면서 하부에 21개의 서브시스템으로 구성된 전체 애플리케이션 구성도는 다음과 같다(그림 2-2 참조).

〔그림 2-1〕 보건복지통계정보시스템 목표시스템 구성도



[그림 2-2] 애플리케이션 전체 구성도



2. 보건복지통계정보시스템 구축 및 운영

보건복지통계정보시스템(이하 정보시스템으로 칭함)은 2010년 1단계 구축을 시작으로 하여 2013년까지 4차에 걸쳐 진행하였다. 2009년은 보건복지통계DB 구축을 위한 정보화전략계획 수립과 더불어 ISP 결과에서 도출된 정보시스템 구축의 기본적인 인프라 환경을 조성하는 해였다. 정보시스템의 본격적인 구축은 ISP 실행계획을 도출한 후, 2010년부터 4년간 2013년까지 수행되었다. 즉, 4년 동안 점차적으로 정보시스템의 기능 및 인프라, 애플리케이션 등을 보완·강화, 개선하는 작업을 진행한 것이다.

〈표 2-1〉은 이용자에게 맞춤형 통계서비스를 제공하겠다는 목표를 달성하고자 수행한 연도별 정보시스템 구축 현황을 간단히 정리한 것이다.

〈표 2-1〉 보건복지통계정보시스템 연도별 구축 현황

연도	구축 내용
2010년	- 보건복지통계포털 구축 ◦ 마이크로데이터 제공서비스 ◦ 통계자료 제공 서비스 ◦ 메타정보 제공 서비스 - 마이크로데이터 관리시스템 구축 ◦ 기초정보관리(레이아웃, 부호화지침서, 조사표) - 통계정보 관리시스템 구축 ◦ 통계표 입력관리 ◦ 통계표 기본정보 관리 ◦ 통계목록관리 및 자료현황 관리 - 코드정보 관리시스템 구축 - 메타데이터관리시스템 구축
2011년	- 통계포털 리뉴얼 및 관리자시스템 개발 - 기능 개발 ◦ 사이트 관리, 통합검색 서비스, 웹진서비스 - 데이터마트 구축 ◦ OLAP 비정형통계 서비스를 위한 애플리케이션 도입 ◦ 데이터모델링 및 데이터 추출, RDB 구축

연도	구축 내용
2012년	- 모바일 웹페이지 구축 - 지역통계 서비스를 위한 GIS 기능 개발 - 보건복지통계 웹진 기능 개선 - 통계포털 리뉴얼
2013년	- 통계분류체계 개편에 따른 DB구조 재설계 - 관리자시스템 전면 개편 - 통계포털 이용자영역 고도화 ◦ UI 전면 개편

3. 통계 DB 구축

구축된 데이터베이스에는 통계표 중심의 보고통계 및 조사 통계, 마이크로데이터, 통계관련 보고서 및 정보 등이 있다.

가. 통계표 중심의 보고통계 DB 구축

보건복지관련 통계 수집은 보건복지통계포털 서비스를 위해 2010년 정보시스템 구축이 진행되면서 국내뿐 아니라 해외 통계까지 수집하였고 이를 매년 지속적으로 데이터베이스화하고 있다. <표 2-2>는 2010년부터 2013년까지 4년간 수집한 보건복지관련 보고통계의 누적현황을 나타낸 것인데, 점차적으로 다양한 분야의 통계가 수집된 것을 보여주고 있다. <표 2-2>에서 보여 주고 있듯이 복지통계는 2013년 기준으로 2,601개의 보고통계를 수집한 상태이며, 타 분야와 비교했을 때 상당히 높은 수집률을 보인다. 2011년에 신규로 수집한 통계는 연구원이 발간하고 있는 통계 중심의 Issue & Focus에 수록된 통계가 데이터베이스화되었으며, 2012년에는 통일을 준비하는 정부나 국민의 복한에 대한 관심이 증

가하면서 북한관련 통계의 수요에 빠르게 대응하기 위해 국내·외 북한통계 수집을 시작하였다. 국제사회와 고립되어 있는 폐쇄적인 북한의 특성으로 인해 보건복지와 관련된 통계를 수집하는 것은 한계가 있었고, 국제기구에 공개되고 있는 북한통계 역시 북한의 공식적 통계보다는 비공식적 통계가 많았다. 그럼에도 불구하고 수집할 수 있는 북한통계를 데이터베이스화하여 서비스하였고, 통계표에 대한 주석을 통해 통계해석의 오류를 줄이는 데 노력하였다. 한편, 빈곤 및 불평등의 실태를 파악할 수 있는 빈곤통계도 2012년부터 새로 수집하였다.

〈표 2-2〉 보건복지관련 통계(조사통계 제외) DB수집 누적현황

(단위: 건)

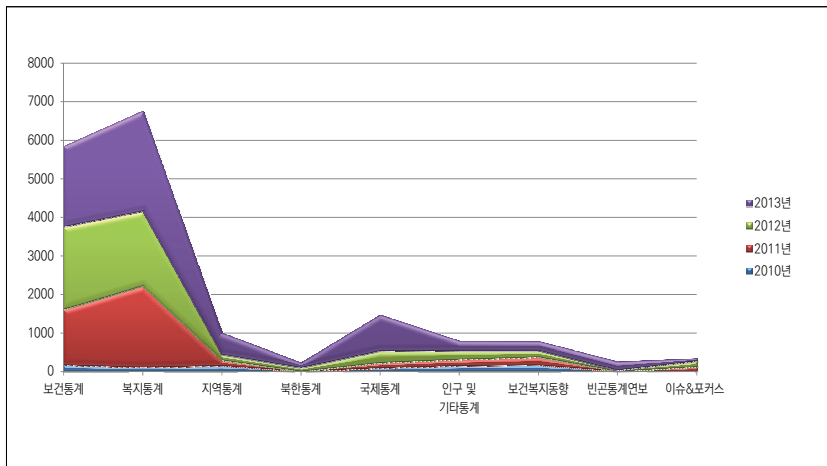
구분	2010년	2011년	2012년	2013년
보건통계	165	1,440	2,143	2,087
복지통계	107	2,112	1,937	2,601
지역통계	146	146	145	564
북한통계	-	-	96	133
국제통계	90	135	308	935
인구 및 기타통계	147	162	234	250
보건복지동향	165	207	172	240
빈곤통계연보	-	-	33	226
Issue&Focus	-	119	153	70
계	820	4,321	5,221	7,106

〈표 2-2〉는 누적현황임에도 불구하고 보건 및 복지통계의 누적건수가 감소하는 해가 있었다. 즉, 보건통계는 2013년에 전년대비 통계 건수가 감소하였고, 복지통계는 2012년에 전년대비 감소하였다. 이유는 중복 또는 유사통계의 삭제 및 정리로 인하여 통계 수가 감소한 것으로 파악되었

다. 한편, 국제통계는 경제협력개발기구(OECD)에서 생산되는 통계 위주의 수집에서, 국제기구 및 국가 단위로 수집처를 확대함에 따라 2013년 수집률이 상당히 높아졌다.

[그림 2-3] 연도별 보건복지관련 통계(조사통계 제외)의 분야별 DB구축 분포도

(단위: 건)



나. 통계표 중심의 조사통계 DB구축

조사통계의 DB구축은 연구원이 직접 수행한 조사사업에서 생산된 통계 중심으로 데이터베이스화하였으며, 구축된 조사통계는 <표 2-3에 제시하였다. 통계청 국가통계포털(KOSIS)은 조사사업보고서에 수록된 통계의 일부를 서비스하고 있으므로 국가통계포털과 차별을 두기 위하여 조사사업보고서에 수록된 모든 통계를 디지털화한 것이 특징이다.

〈표 2-3〉 통계표 중심의 조사통계 DB구축 현황

(단위: 건)

구분	조사통계 명	2012년	2013년 (누적)
보건 통계	국민보건의료 실태조사	25	25
	환자조사	466	455
	한국의료패널조사	105	105
	의약품 소비량 및 판매액 통계조사	106	106
	영아모성사망조사	114	114
	청소년 건강행태 온라인 조사		84
	소계	816	889
복지 통계	국가유공자 생활실태조사	492	492
	국민기초생활보장 수급자 현황	38	38
	결혼 및 출산 동향 조사	569	894
	노인생활 실태 및 복지욕구 조사	779	700
	노인 장기요양 보호 욕구 실태조사	61	61
	노인학대 실태조사	78	78
	능동적 복지확충을 위한 복지실태 및 복지욕구 조사	454	454
	보육실태조사	438	438
	불임 및 불임 관련 의료이용 실태	34	34
	장애인 실태조사	752	1,287
	저소득층 자활사업 실태조사	353	353
	차상위 계층 실태조사	794	794
	출산력 및 가족 보건복지 실태조사	721	728
	전국 다문화 가족 실태조사	554	554
	한국복지패널 조사	635	635
	베이비부머의 생활실태 및 복지욕구	211	211
	소계	6,963	7,751

자료: 오미에 외, 「2013 보건복지통계정보시스템 구축 및 운영」, 재인용.

다. 마이크로데이터 DB구축

통계포털의 가장 차별화된 영역은 바로 조사사업 수행을 통해 생산된 마이크로데이터 DB구축이다. 마이크로데이터 DB구축은 18대 정부가 추구하는 정부 3.0의 공공정보 개방 및 공유를 과거부터 실천한 사례라 할 수 있다. 마이크로데이터 DB구축은 보건복지와 관련된 조사데이터를 공개함으로써 일반인들에게 데이터의 접근성을 향상하고, 활용성을 높이며 데이터의 체계적 관리를 가능케 하였다. 데이터가 공개되고 있는 조사사업은 보건 2종, 복지부문 11종으로서 “출산력 및 가족 보건복지 실태조사”와 같이 주기성을 가진 사업은 각 주기에 생산된 마이크로데이터가 대부분 제공된다.

〈표 2-4〉의 각 마이크로데이터와 관련된 부가자료 즉, 조사표 및 코딩 지침서, 레이아웃 등도 함께 데이터베이스화하여 일반인이 쉽게 이용할 수 있도록 하였다. 〈표 2-4〉에서 보여 주듯이 가장 많은 데이터를 보유한 조사사업은 환자조사로서 과거 1999년부터 최근 2015년까지 생산된 모든 조사데이터를 필요에 따라서 누구라도 이용할 수 있도록 서비스하고 있다.

라. 기타 통계관련 DB구축

정보시스템은 통계와 마이크로데이터뿐 아니라 이와 연관된 분석보고서나 자료집, 또는 통계용어, 통계뉴스 등 다양한 부가적 콘텐츠를 주기적으로 수집하여 데이터베이스화하였다(표 2-5 참조).

제2절 보건복지데이터포털 운영 현황

보건복지통계포털은 보건복지 분야의 각 사이트 및 보고서에 산재되어 있는 보건 및 복지관련 통계를 집대성해 보여 주기 위한 목적으로 구축·운영되었다. 그러나 2014년에는 통계청에서 제공하는 국가통계포털(KOSIS)의 막대한 정보와 차별성을 두고자, 승인된 통계의 가치 있는 마이크로데이터를 집적화하여 체계적으로 집중 관리 할 조직 및 사이트인 보건복지데이터포털로 개편을 진행하였다.

보건복지데이터포털의 가장 차별화된 영역은 바로 조사사업 수행을 통해 생산된 마이크로데이터 DB구축이다. 마이크로데이터 DB구축은 정부 3.0의 공공정보 개방 및 공유를 과거부터 실천하고 있는 사례라고 할 수 있다. 따라서 보건복지데이터포털은 보건복지와 관련된 데이터에 대한 정보를 공개하고 제공함으로써 일반인들의 데이터의 접근성을 향상하고, 활용성을 높이며, 데이터를 체계적으로 관리하고자 하였다.

통계표 중심의 보건복지통계포털에서 데이터 중심의 보건복지데이터포털로 개편작업을 실시한 내용을 개편 전과 개편 후의 메인페이지 화면으로 보여 줌으로써 어느 부분이 달라졌는지를 한눈에 알아볼 수 있도록 하였다. 2014년과 2015년의 주요 개편 내용은 연구 DB, Data Bank (Micro & Macro data), 해외자료(해외 Micro & Macro data, 국제기구통계), 인포그래픽스와 공간정보통계이다.

1. 연구 DB

연구 DB는 한국보건사회연구원에서 수행한 연구과제 정보를 이와 관련된 Micro & Macro 데이터의 메타정보와 연계하여 제공함으로써, 수

요자에게 더 풍부하고 다양한 정보를 제공하는 데 주된 목적이 있다. 연구과제는 기본연구(연구보고서)와 수탁연구(정책보고서)를 포함하고 있고, 관련 정보를 연구주제별, 생애주기별, 명칭별로 구분하였다.

연구주제별 분류체계는 트리 형태로 구조화되어 있으며, 대분류 체계는 인구/가족(인구변화, 생애주기, 가족), 보건의료(건강보장, 건강증진, 보건산업, 보건일반), 사회복지(사회보험, 공공부조, 사회서비스, 복지행정, 민간복지, 복지일반), 지역사회(농어촌, 지역복지), 사회통합/행복/삶의 질, 소득/재산/소비, 빈곤/불평등, 조세/재정, 조사/통계/정보, 통일/보훈, 복지국가비교, 미래사회로 구성되어 있다. 생애주기별 분류체계는 영유아, 아동/청소년, 청년, 베이비부머, 노인, 장애인, 여성으로 구분하였고, 명칭별 분류체계는 ㄱ~ㅎ(자음), 0~9(숫자), A~Z(영문)로 구분하였다.

[그림 2-4] 보건복지데이터포털 기본구조 및 주요내용



〈표 2-4〉 2014년 보건복지데이터포털 연구 주제별 분류체계

대분류	중분류	세분류	
(A) 인구/가족	(1) 인구변화	(1) 저출산	(2) 고령화
		(3) 출생/사망	(4) 결혼/이혼
		(5) 이민/입양	
	(2) 생애주기	(1) 영유아	(2) 아동/청소년
		(3) 청년	(4) 베이비부머
		(5) 노인	(6) 장애인
		(7) 여성	
	(3) 가족	(1) 한부모	(2) 다문화
(B) 보건의료	(1) 건강보장	(1) 의료안전망	(2) 의료자원
	(2) 건강증진	(1) 생활습관병	(2) 정신건강
		(3) 만성질환	
	(3) 보건산업	(1) 의료산업	(2) 의약품(화장품)산업
		(3) 식품산업	
	(4) 보건일반		
(C) 사회복지	(1) 사회보험	(1) 국민연금/특수직역연금/퇴직연금/개인연금	
		(2) 건강보험	(3) 노인장기요양보험
		(4) 고용보험	(5) 산재보험
	(2) 공공부조	(1) 기초생활보장	(2) 근로빈곤/자활
		(3) 차상위지원	(4) 기초연금/장애인연금
	(3) 사회서비스	(1) 보육/교육	(2) 돌봄(노인/장애인)
		(3) 주거/에너지/환경	(4) 고용/일자리
		(5) 문화	
	(4) 복지행정	(1) 전달체계	(2) 바우처
	(5) 민간복지	(1) 복지자원	(2) 사회참여/자원봉사
		(3) 사회공헌/기부	
	(6) 복지일반		
(D) 지역사회	(1) 농어촌		
	(2) 지역복지		
(E) 사회통합/행복/삶의 질			

대분류	중분류	세분류
(F) 소득/재산/소비		
(G) 빈곤/불평등		
(H) 조세/재정		
(I) 조사/통계/정보		
(J) 통일/보훈		
(K) 복지국가비교		
(L) 미래사회		

〈사회보장통계 관련 주요 개편 내용〉

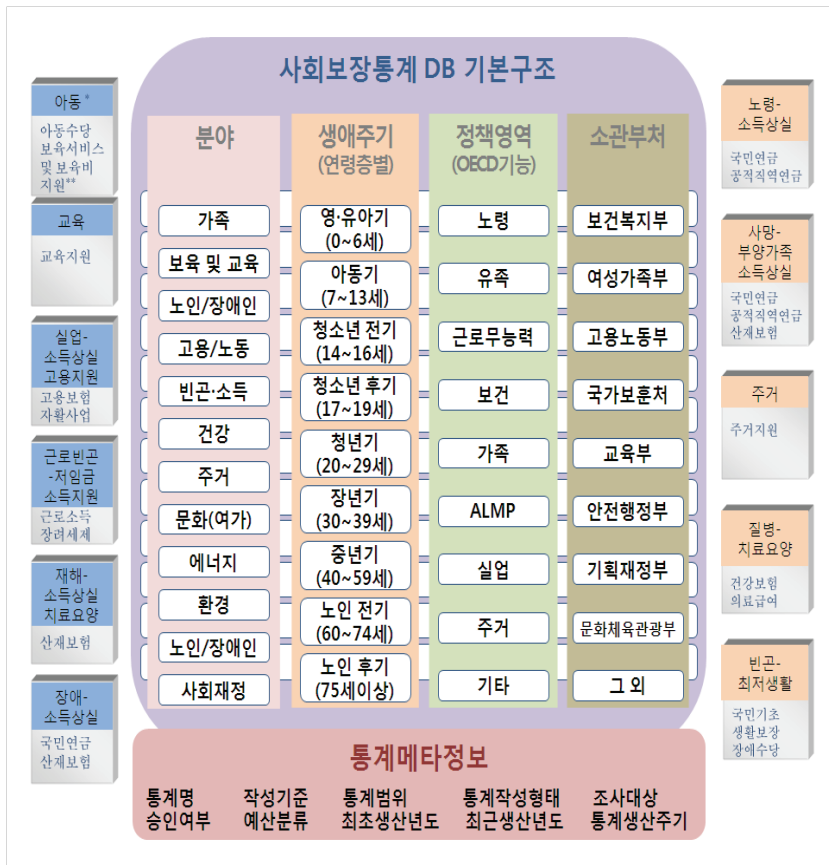
사회보장은 출산, 양육, 노령, 장애, 질병, 빈곤, 사망 등의 사회적 위험으로부터 국민을 보호하고 국민 삶의 질을 향상하는 데 필요한 소득·서비스를 보장하는 사회보험, 공공부조, 사회서비스를 의미한다. 이에 따라 우리나라의 맞춤형 복지사회 구현을 뒷받침하기 위해 근거 중심의 정책 평가 및 설계가 필요하고, 사회보장의 실태를 진단할 수 있는 통계의 통합적 생산 및 관리의 중요성이 증대되고 있다.

이에 따라 사회보장위원회에서는 효과적인 사회보장정책의 수립을 지원하기 위해 체계적인 사회보장통계 조사·분석 및 관리를 하고 있으며, 사회보장통계는 국민의 생애주기 전반에 걸쳐 발생하는 각종 사회적 위험의 정도와 이에 대응하는 정부 정책의 내용과 그 성과를 주요 내용으로 하고 있다. 사회보장통계의 지표선정은 사회보장위원회를 통해 검토, 확정되었으며, 사회보장통계 총괄·관리는 사회보장위원회 사무국이 담당하고 있다.

보건복지데이터포털의 사회보장통계에서는 사회보장과 관련하여 각 기관에서 제출한 사회보장통계 중 '사회보장의 현재 발전 상황이나 주요 정책성과를 직관적으로 판단하기 위해 반드시 필요한 지표'를 제시하고

자 하였다. 지표는 생애주기 단계에서 발생하는 사회적 위험(육구)과 정책 대응의 내용에 따라 분야별, 생애주기별, 정책영역별, 소관부처별로 제시하고 모두 253개의 지표를 대표지표와 관련된 연관지표로 구분하여 각 지표별로 정의, 측정 산식, 관련 통계정보 등을 보여 준다. 세부적으로 가족, 아동·노인·장애인, 보육 및 교육, 근로, 소득보장과 빈곤, 건강, 주거, 에너지, 환경, 문화, 사회재정으로 구분되어 있다.

[그림 2-5] 사회보장통계 DB 구축안



2. Data Bank(Micro & Macro Data)

Data Bank에서는 한국보건사회연구원에서 그동안 수행한 연구과제에서 조사·수집된 Micro & Macro 데이터의 메타정보와 조사표 및 원자료 등을 제공한다. 연구과제로부터 생산된 다양한 데이터들을 연구주제별, 데이터유형별로 구분하였으며, Data Summary에서는 여기에 수록된 다양한 주요 마이크로데이터의 메타정보를 요약하여 여러 내용을 쉽게 비교할 수 있도록 제공하였다.

연구주제별 Micro & Macro 데이터는 한국보건사회연구원에서 그동안 수행한 연구과제 정보를 제공하는 연구 DB의 연구주제별 분류체계 및 개별 연구과제 정보와 연계하여 조사목적, 조사방법, 조사대상, 표본설계, 통계승인 여부, 데이터 제공 여부 등의 풍부한 정보를 제공하고자 하였다.

데이터유형별 Micro & Macro 데이터는 Micro data(양적), Micro data(질적), Macro data로 구성되며 데이터 유형별로 제공하고 있는 데이터의 목록을 한눈에 확인할 수 있다. Micro data(양적, Quantitative)는 주로 변수에 대한 수량적 자료를 산출할 수 있는 설문지, 면접용 조사표 등의 방법으로 수집된 데이터이고, Micro data(질적, Qualitative)는 주로 연구대상과 긴밀하게 관계를 유지할 수 있는 참여관찰, 개방형 비조직적 면접, 사적 기록물, 포커스그룹 인터뷰 등의 방법으로 수집된 데이터이다. Macro data에서는 국가, 지역 등을 단위로 하여 생산된 다양한 통계정보를 특정 기준에 따라 수집하여 축적한 데이터를 제공한다.

〈표 2-5〉 보건복지데이터포털 Data Summary(주제별¹⁾)

구 분	일반적 특성	보건 및 의료	주거	가족, 친구, 이웃관계	임신/출산 /피임/ 모자보건	결혼생활 및 준비/ 가족 주기	일상생활 지원/ 활동제한	보육 및 교육	취업 및 직업생활
장애인실태조사									
가구 및 장애판별조사	★★		★						
장애인 심층조사	★★★★	★★	★		★	★	★★	★★	★★★★
사회복지시설 조사(행복e음)	★								
환자조사									
기관조사	★	★★★★							
외래환자	★	★★★★							
퇴원환자	★	★★★★					★		
한국복지패널									
가구조사	★	★★	★★	★					★★
가구원조사		★★		★	★				★
부가조사 ²⁾									
한국의료패널									
가구조사	★★★	★★★★	★		★★				★★
성인가구원조사		★★★★					★★		★
노인실태조사									
가구조사	★								
노인가구원조사		★★★		★★★			★★★★		★
전국 출산력 및 가족보건·복지실태조사									
가구조사	★★		★		★				
15~49세 기혼여성조사				★★★★	★★★★	★★★		★	★
20~44세 미혼남녀조사				★★	★	★★			★

주: 1) ★ 수준은 정보의 양과 조사 문항 수에 따라 정해졌고, ★ 수가 많을수록 관련성이 많음을 의미함.

2) 한국복지패널은 차수마다 복지인식, 아동, 장애인 부가조사로 구성됨.

〈표 2-6〉 보건복지데이터포털 Data Summary(주제별¹⁾ II)

구 분	사회 및 여가 활동	생활 만족도 및 폭력/ 차별 경험	경제 상태	보험, 의료 관련 비용 및 보장	사회보장, 복지 서비스 및 인식 등	여성 장애인	장애인 보조 기구	장애인 부가조사	아동 부가조사
장애인실태조사									
가구 및 장애판별 조사	.	★	★	.	.	.	.	.	.
장애인 심층조사	★	★	★	.	★	★	★	.	.
사회복지시설 조사(행복e음)	.	.	.	.	.	.	.	.	.
환자조사									
기관조사	.	.	.	.	.	.	.	.	.
외래환자	.	.	.	.	.	.	.	.	.
퇴원환자	.	.	.	.	.	.	.	.	.
한국복지패널									
가구조사	.	★	★★★★	★	★★	.	.	.	.
가구원조사	★	★	.	★	.	.	.	.	.
부가조사 ²⁾	.	.	.	.	★★★	.	.	★★★★	★★★
한국의료패널									
가구조사	.	.	★★	★★★	.	.	.	.	.
성인가구원조사	.	★	.	★	★★	.	.	.	.
노인실태조사									
가구조사	.	.	★	.	.	.	.	.	.
노인가구원조사	★★	★★	★★★	.	★★	.	.	.	.
전국 출산력 및 가족보건·복지실태조사									
가구조사	.	.	★	.	.	.	.	.	.
15~49세 기혼여성조사	.	.	.	.	★★	.	.	.	.
20~44세 미혼남녀조사	.	.	.	.	★	.	.	.	.

주: 1) ★ 수준은 정보의 양과 조사 문항 수에 따라 정해졌고, ★ 수가 많을수록 관련성이 많음을 의미함.

2) 한국복지패널은 차수마다 복지인식, 아동, 장애인 부가조사로 구성됨.

〈표 2-7〉 보건복지데이터포털 Data Summary(조사방법)

구 분	면접조사	자가기입 조사	등록 데이터
장애인실태조사			
가구 및 장애판별 조사	√	.	.
장애인 심층조사	√	.	.
사회복지시설 조사(행복e음)	.	.	√
환자조사			
기관조사	.	.	√
외래환자	.	.	√
퇴원환자	.	.	√
한국복지패널			
가구조사	√	.	.
가구원조사	√	.	.
부가조사	√	.	.
한국의료패널			
가구조사	√	.	√
성인가구원조사	√	.	√
노인실태조사			
가구조사	√	.	.
노인가구원조사	√	.	.
전국 출산력 및 가족보건·복지실태조사			
가구조사	√	.	.
15~49세 기혼여성조사	√	.	.
20~44세 미혼남녀조사	√	.	.

〈표 2-8〉 보건복지데이터포털 Data Summary(조사범위)

구 분	전국 대표 표본	전수
장애인실태조사		
가구 및 장애판별 조사	●	·
장애인 심층조사	●	·
사회복지시설 조사(행복e음)	·	●
환자조사		
기관조사 ¹⁾	●	·
외래환자	●	·
퇴원환자	●	·
한국복지패널 ²⁾		
가구조사	●	·
가구원조사	●	·
부가조사	●	·
한국의료패널		
가구조사	●	·
성인가구원조사	●	·
노인실태조사		
가구조사	●	·
노인가구원조사	●	·
전국 출산력 및 가족보건·복지실태조사		
가구조사	●	·
15~49세 기혼여성조사	●	·
20~44세 미혼남녀조사	●	·

주: 1) 환자조사의 기관조사는 전수층과 표본층으로 나뉨. 전수층에는 비중이 큰 의료기관이나 조사의 관리 또는 특성상 꼭 조사되어야 할 7개 종류의 기관(종합병원, 일반병원, 치과병원, 한방병원, 보건의료원, 보건소, 조산원)이 속함.

2) 한국복지패널은 복지욕구를 더 효과적으로 파악하기 위하여 저소득층을 과대표집 하였음.

〈표 2-9〉 보건복지데이터포털 Data Summary(조사단위)

구 분	개인	가구	기관
장애인실태조사			
가구 및 장애판별 조사	.	●	.
장애인 심층조사	●	.	.
사회복지시설 조사(행복e음)	●	.	.
환자조사			
기관조사	.	.	●
외래환자	●	.	.
퇴원환자	●	.	.
한국복지패널			
가구조사	.	●	.
가구원조사	●	.	.
부가조사	●	.	.
한국의료패널			
가구조사	.	●	.
성인가구원조사	●	.	.
노인실태조사			
가구조사	.	●	.
노인가구원조사	●	.	.
전국 출산력및 가족보건·복지실태조사			
가구조사	.	●	.
15~49세 기혼여성조사	●	.	.
20~44세 미혼남녀조사	●	.	.

〈표 2-10〉 보건복지데이터포털 Data Summary(조사대상)

구 분	영·유아	아동/청소년	청년	중장년	노인	장애인
장애인실태조사						
가구 및 장애판별 조사	●	●	●	●	●	●
장애인 심층조사	●	●	●	●	●	●
사회복지시설 조사(행복e음)	●	●	●	●	●	●
환자조사						
기관조사						
외래환자	●	●	●	●	●	●
퇴원환자	●	●	●	●	●	●
한국복지패널						
가구조사	●	●	●	●	●	●
가구원조사		●	●	●	●	●
부가조사		●	●			●
한국의료패널						
가구조사	●	●	●	●	●	●
성인가구원조사			●	●	●	●
노인실태조사						
가구조사	●	●	●	●	●	●
노인가구원조사					●	●
전국 출산력 및 가족보건·복지실태조사						
가구조사	●	●	●	●	●	●
15~49세 기혼여성조사	●	●	●	●	●	
20~44세 미혼남녀조사			●	●		

〈표 2-11〉 보건복지데이터포털 Data Summary(조사주기¹⁾)

구 분	1994년 이전	1995~ 1999년	2000~ 2004년	2005~ 2009년	2010~ 2015년	총 조사 횟수
장애인 실태조사	1980(1) 1985(2) 1990(3)	1995(4)	2000(5)	2005(6) 2008(7)	2011(8) 2014(9)	9회
환자조사	1953(1) 1985(8) 1986(9) 1988(10) 1990(11) 1992(12) 1994(13)	1996(14) 1999(15)	2002(16)	2005(17) 2008(18) 2009(19)	2010(20) 2011(21) 2012(22) 2013(23) 2014(24)	24회
한국복지패널				2006(1) 2007(2) 2008(3) 2009(4)	2010(5) 2011(6) 2012(7) 2013(8) 2014(9) 2015(10)	10회
한국의료패널				2008 2009	2010 2011 2012 2013 2014 2015	8회
노인실태조사	1994(1)	1998(2)	2004(3)	2008(4)	2011(5) 2014(6)	6회
전국 출산력 및 가족보건·복지 실태조사	1964(1) 1991 1994	1997	2000 2003	2006 2009	2012(22) 2015(23)	23회

주: 1) 괄호 안의 숫자는 조사 횟수를 의미함. 조사 횟수가 부정확한 연도는 괄호가 없음.

〈표 2-12〉 보건복지데이터포털 Data Summary(메타데이터)

구 분	보고서	설문지	코드북	매뉴얼
장애인실태조사 (2014)				-
환자조사 (2014)				-
한국복지패널 (2015)				
한국의료패널 (2013)				-
노인실태조사 (2014)				
전국 출산력 및 가족 보건·복지 실태조사 (2015)				-

〈표 2-13〉 보건복지데이터포털 Data Summary(2016년 12월 1일 기준)

구 분	데이터셋	개체수 ¹⁾	변수수	종단 연구	데이터셋 버전날짜	현재 제공되는 데이터 조사연도
장애인실태조사(2014)	1	6,824	986		2016. 5. 2	2000 2005 2008 2011 2014
환자조사(2014)	3					1996 1999
기관조사	1	5,545	215			2002 2005
외래환자	1	798,752	15			2008 2009
퇴원환자	1	617,194	26		2016. 5. 3	2010 2011 2012 2013 2014

구 분	데이터셋	개체수 ¹⁾	변수수	종단 연구	데이터셋 버전날짜	현재 제공되는 데이터 조사연도
한국복지패널(2015)				√	2016. 4. 7.	2006
가구조사	1	6,914	1,082			2007
가구원조사	1	13,647	233			2008
부가조사	1	471	283			2009
						2010
						2011
						2012
						2013
						2014
						2015
한국의료패널(2013)				√	2016. 7. 12.	2008
가구조사	2	5,200	78			2009
성인가구원조사	2	14,839	99			2010
						2011
						2012
						2013
노인실태조사(2014)	1	10,451	720		2015. 8. 31.	1994
						1998
						2004
						2008
						2011
						2014
전국 출산력 및 가족보건·복지실태조사(2015)					2016. 7. 25.	1991
15~49세 기혼여성조사	1	11,009	1,097			1994
						1997
						2000
						2003
						2006
						2009
						2012
						2015

주: 1) 개체 수에는 가중치를 반영하지 않음.

3. 해외자료

보건복지데이터포털에서는 해외 Micro & Macro data 부분을 새롭게 추가하고, 기존의 국제통계 국제기구통계로 정리·보완하였다. 해외 Micro & Macro 데이터 메타정보를 제공하고 링크를 해 줌으로써 유용한 해외자료에 대한 접근 가능성을 높이는 데 그 목적이 있다.

해외 연구기관 및 국제기구 등과 연계하여 마이크로데이터(패널 등 조사자료)와 매크로데이터(국가별 통계지표 등)에 대한 링크 및 데이터 이용방법에 대해 자세히 설명하여 이용자들이 원하는 해외자료를 손쉽게 찾고 이용할 수 있도록 하였다.

Micro 데이터 및 Macro 데이터 관련 메타정보 자료수집 및 체계화에 대한 개념트는 해외 사례(EUI; European University Institute)²⁾ 연구를 참고하여 구성하였다.

해외자료는 해외 Micro & Macro data와 국제기구통계로 구분하였으며, 해외 Micro & Macro 데이터는 Micro 부분과 Macro 부분으로 나누어 보이도록 하고, 데이터명, 데이터설명, 제공연도, 이용방법, 비고, 홈페이지/자료링크로 구성하였다. 국제기구통계는 기존의 통계표를 지표명으로 변경하여 지표[예시: 경제협력개발기구(OECD)에 제출하는 통계인 기대수명]의 개요, 출처, 담당부서, 작성주기, 작성체계, 용어정의, 비교, 주요통계표를 제시하여 관련 메타정보를 제공하고자 하였다.

2) 유럽연합대학교 EUI는 국제 대학원 및 연구기관이며, 1972년 유럽 6개 나라(벨기에, 프랑스, 독일, 이탈리아, 룩셈부르크, 네덜란드)에서 박사과정의 학생(연구원)들에게 향상된 교육을 제공하고 경제, 역사, 정치와 사회과학 관련 법에 대한 연구를 수행하기 위해 설립되었음. 현재 EUI 데이터 포털은 75개의 주요 통계 DB(Micro data-32개, Macro data-43개)의 정보를 제공하고 있으며 EUI회원들에게만 EUI에서 구축한 데이터의 접근을 허용하고 있음.

〈표 2-14〉 보건복지데이터포털 해외 Micro & Macro data 목록

Micro Data	Macro Data
Consumer Expenditure Survey(USA)	Comparative Political Data Sets
Current Population Survey(USA)	Demographic Yearbook(UN)
DNB Household Survey(Netherlands)	Direction of Trade Statistics(IMF)
EU-Labour Force Survey	European Regional Database
EU Statistics on Income & Living Conditions	EU Open Data Portal
European Social Survey(ESS)	Eurostat Data / Eurostat Yearbook
European Values Study	Gender data - United Nations
G-SOEP - German Socio-Economic Panel(DIW Berlin)	GENESIS - German Official Statistics Portal
ICPSR data archive	Government Finance Statistics(IMF)
LIS data on income and inequality	Human Development Indicators & Index(UN)
National Longitudinal Surveys(USA)	International Debt Statistics(World Bank)
Panel Study of Income Dynamics(USA)	International Labour Organization data
Survey of Consumer Finances(USA)	International Monetary Fund databases
SHARE(Data on health and retirement in Europe)	Nordic Statistics - Statistik utan granser
Survey on Household Income and Wealth(Italy)	OECD data & documentation
Survey of Income and Program Participation(USA)	Social Security Data(ISSA)
Swedish Socio-economic Data	United Nations data
UK Data Service and Archive	UNICEF ChildInfo Data Portal
U.S. Census	World Bank data portal
World Values Survey	World Development Indicators(World Bank)

4. 인포그래픽스와 공간정보통계

인포그래픽스와 공간정보통계는 2015년의 개편 내용이다. 데이터 시각화 부분인 인포그래픽스는 많은 이용자들이 원했던 콘텐츠로, 일반인들이 이해하기 쉽도록 데이터를 시각적으로 변환하여, 한눈에 쉽게 볼 수 있도록 하면 좋겠다는 의견을 반영하여 통계표를 시각화하여 제공하였고, 주요실태조사와 보건복지데이터포털의 홍보를 중심으로 하여 인포그래픽스를 제작하였다. 인포그래픽스 및 통계 시각화는 이용자들에게 복잡하고 어려운 정보를 쉽고 간결한 구조로 시각화하여 전달하는 데 목적이 있다. 따라서 이는 보건복지분야 관련 통계 자료에 대한 국민 이해도 및 수용도를 높일 수 있도록 최근 경향을 반영한 홍보 콘텐츠라고 할 수 있다. 메인홈페이지도 주요 콘텐츠 중심으로 공간을 구성하되, 데이터 시각화인 인포그래픽스를 메인홈페이지 중앙에 배치하여 이용자들의 흥미를 끌 수 있도록 하였다.

공간정보통계는 시도통계와 시군구통계로 나누어져 있으며, 세종특별자치시에 대한 지역 정보를 구분하여 제공할 수 있도록 지도 플랫폼 기능을 향상하였다. 공간정보통계는 과거의 ‘지역통계’를 고도화한 콘텐츠로, 시도 및 시군구 통계를 시각적으로 표현하였을 뿐만 아니라, 전에는 없었던 맞춤형 공간정보통계 시각화 서비스를 제공하였다. 이용자가 생산한 시도 또는 시군구 데이터를 공간정보통계의 ‘맞춤형 통계 시각화’에 올려 통계 수치에 따라 구분할 수 있는 지도차트 이미지 파일을 제공하며 이를 저장하거나 인쇄할 수 있도록 구현하였다.

2016년에는 연구원에서 생산한 2015년 전국 출산력 및 가족보건·복지 실태조사 결과를 인포그래픽스로 제작하였으며, 공간정보통계 영역도 맵차트 분할 모드를 제공하여 고도화 작업을 진행하였다.

제 3 장

보건복지데이터포털 개편 및 신규 콘텐츠 구축

제1절 보건복지데이터포털 개편 내용

제2절 신규 콘텐츠

3

보건복지데이터포털 개편 및 신규 콘텐츠 구축

제1절 보건복지데이터포털 개편 내용

2014년 보건복지데이터포털의 개편에서 연구 DB, Data Bank (Micro & Macro data), 해외자료(해외 Micro & Macro data, 국제기구통계)로 나누어, 데이터 중심으로 큰 틀을 구축했다면, 2015년 개편에서는 이용자 중심의 포털이 될 수 있도록 인포그래픽스와 공간정보통계로 고도화 작업을 진행하였다. 2016년의 주요 개편방향은 Recommendation system 구축과 Deposit data section 개발이다. Recommendation system 도입은 이용자 맞춤형 서비스를 제공할 수 있게 한다. 이용자 데이터베이스와 보건복지데이터포털의 데이터베이스를 결합하여 모형을 구축하며, 이용자가 원하는 정보를 Recommendation system을 통해 제공해 줌으로써 정보 제공의 다양성을 확보할 수 있다. Deposit data는 데이터 위탁관리 및 공유 서비스를 의미한다. 보건복지데이터포털에서는 보건복지 영역 또는 관련 분야에서 수행된 모든 연구과제에서 생산, 활용된 다양한 데이터를 위탁관리 하고 이와 더불어 다른 연구자의 데이터를 공유할 수 있도록 지원하고자, Deposit data section을 개발하였다. 보건복지 분야 연구자들의 연구 성과 및 제공된 데이터는 한국보건사회연구원에서 수행한 다양한 연구의 과제 및 데이터와 동일하게 등록, 관리되며, 다른 연구자들의 연구결과를 인용하고 데이터를 분석에 활용함으로써 연구자들의 연구 업적이 제고될 수 있도록 지원해 주는 역할을 한다.

■ 콘텐츠

- 3) 본 내용 중 '수정(개선 전)' 부분은 [오미애 외(2015). 2015년 보건복지통계정보 통합 관리 및 운영, 한국보건사회연구원의 일부 내용을 요약발췌(재인용)한 것임.

화면명	메인	구분	수정(개선 전)
			

■ 메인 콘텐츠 구성

- 주요 콘텐츠인 ‘신규 연구 DB’, ‘신규 Micro & Macro’, ‘해외통계’, ‘공지사항’ 중심으로 공간을 크게 나누어 시각적으로 시원한 모습을 보여 줌.
- ‘신규 연구 DB’와 ‘신규 Micro & Macro’는 탭방식으로 구현하여, 한정된 메인페이지를 효율적으로 사용함.
- 신규 연구 게시물의 내용을 일부 보이도록 하여, 내용에 대해 미리 확인할 수 있도록 표현함.

○ 사용자가 많이 조회하는 자료를 메인 화면에 노출함으로써 콘텐츠 접근성을 높임.

화면명	메인-인기자료	구분	수정(개선 후)
-----	---------	----	----------

화면명

메인-인기자료

구분

수정(개선 후)

[illegible]

■ 수행 로직

- 조회 수가 많은 자료 5개 중에서 3개를 무작위로 출력하여, 같은 콘텐츠가 계속해서 노출되는 현상을 예방함.
- 사이트 운영상 필요한 경우, 관리자가 직접 인기자료 콘텐츠를 지정할 수 있음.
- 관리자가 직접 지정한 인기자료는 각 게시물마다 70% 확률로 메인 화면에 노출됨. 즉, 메인 화면을 10번 접속할 경우 7번 정도 관리자가 지정한 인기자료가 메인 화면에 노출됨. 70% 확률로 노출하는 이유는, 노출되는 콘텐츠가 고정되는 현상을 방지하기 위한 것임.
- 만약 관리자가 2개의 추천 콘텐츠(A, B)를 지정하면, 총 3개의 콘텐츠가 들어갈 수 있는 메인 화면 공간에, 관리자 지정 콘텐츠 A - 70%, B - 70% 확률로 노출 여부를 결정한 후에 남은 1개의 콘텐츠 자리는 조회 수가 많은 자료 5개 중에서 무작위로 선정함.

화면명	연구DB - 주제별	구분	수정(개선 전)
------------	-------------------	-----------	-----------------

로그인 마이페이지 회원가입 사이트맵 서비스문의 검색어를 입력해주세요 🔍



[연구DB](#)
[Data Bank](#)
[사회보장통계](#)
[공간정보통계](#)
[해외자료](#)
[인포그래픽스](#)
[정보마당](#)

연구 DB

주제별

- 연구과제
 - 연구과제
 - 정책화
 - 홍성/사업
 - 결과분석
 - 이벤트/심사
- 생애주기
 - 가정
 - 노년
- 사회안전망
 - 사회복지
 - 지역사회
 - 사회통합/참여/삶의질
 - 소득/재산/소비
 - 빈곤/불평등
 - 조사/통계/정보
 - 통일/보훈
 - 복지국가비전
 - 미래사회

생애주기별

영향범위

주제별

🔍

번호	연구과제명(대여명)	발행년도
53	2012년도 난임부부 지원사업 결과분석 및 평가 (난임부부 지원 사업 평가를 위한 조사·제외·수정 연..	2013
52	가임기 여성 임신 전 출산건강 관리지원 방안 연구(출몰보고서)	2013
51	고령 임신부의 산전 질환 및 출산결과 통합분석과 정책적 함의 (교원(회)교원) 임신부 정의 및 정책 수..	2013
50	고령 임신부의 산전 질환 및 출산결과 통합분석과 정책적 함의	2013
49	산업단지 어린이집 확대 방안 연구 (산업단지근로자를 위한 어린이집 설치여건 조사(연립조사))	2013
48	산업단지 어린이집 확대 방안 연구 (산업단지근로자를 위한 어린이집 설치여건 조사(실용조사))	2013
47	임신 전 출산건강 관리 실태 및 지원 방안 (임신준비 및 출산에 대한 조사)	2013
46	임신 전 출산건강 관리 실태 및 지원 방안 (최근 출산여성 및 전문가 설문인터뷰)	2013
45	저출산 고성장 시대의 한국 가족주의에 대한 진단과 정책적 함의	2013
44	활착한 임신출산을 위한 기저귀조제품 지원 방안 연구 (영아 자녀의 소유 및 기저귀 이용 실태조사)	2013

<< < **1** 2 3 4 5 6 > >>

개인정보처리방침 | 이용약관 | RSS

30547 서울특별시지식시설팀(대표: 370 서울국제특수교육지원센터) | TEL : 044-267-8007 / FAX : 044-267-8009 / EMAIL : micro@khanacsr.com

Copyright © by KHANACSR All rights Reserved.

본 홈페이지에서 게시된 이메일주소 등 자동으로 수집되는 것을 거부하며, 이를 위반시 관련법에 따라 처벌받을 수 있습니다.

■ 화면 구성

- 주제별 트리구조는 분류체계에 따라 구성됨.
- 주제별 분류체계는 연구과제의 중요단어(keyword)를 고려하여 연구진 회의에서 분류작업을 수행하였으며, 각 연구과제는 하나의 범주에만 포함되지 않고 여러 주제별 트리구조에 포함됨.
- 선택한 연구 주제에 따라 연구과제 목록이 발행 연도 기준으로 정렬되어 출력됨.

화면명	연구DB - 주제별 - 인기자료	구분	수정(개선 후)																																	
 연구DB Data Bank 사회보장통계 공간정보통계 해외자료 인포그래픽스 정보마당 연구 DB 주제별 연구/가주 인기자료 주제별 인기자료 2012년 전국 결혼 및 출산 동향 조사 (2012년 전국 출산력 및... - 가임기 여성 임신 전 출산연령 관리지원 방안 연구(출발보고서) - 임신 전 출산연령 관리 실태 및 지원 방안 (임신준비 및 출산에... 주제별 최신자료 - 고령화 및 생산가능인구 감소에 따른 대응전략 마련 연구 2014년 저출산고령사회정책 성과평가 (저출산고령사회정책 관련 전문가 조사) 2014년 저출산고령사회정책 성과평가 (저출산고령사회정책의 성과에 대한 국민 체감도 조사) - 출산율 반등세로의 전환을 위한 대응전략 마련 연구 - 산후조리원 걱정 운영기준 마련에 관한 연구 (산후조리원 걱정운영 방안에 관한 전문가 조사) - 산후조리원 걱정 운영기준 마련에 관한 연구 (산후조리원 실용면접조사) - 산후조리원 걱정 운영기준 마련에 관한 연구 (산후조리원 실태조사) 2015년 전국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(미혼)) 2015년 전국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(기혼)) 2014년도 난임부부 지원사업 결과분석 및 평가 (난임부부 지원사업에 대한 만족도 및 요구도 조사) 연구과제명 검색 <table border="1"> <thead> <tr> <th>번호</th> <th>연구과제명(데이터명)</th> <th>발행년도</th> </tr> </thead> <tbody> <tr><td>53</td><td>고령화 및 생산가능인구 감소에 따른 대응전략 마련 연구</td><td>2015</td></tr> <tr><td>52</td><td>2014년 저출산고령사회정책 성과평가 (저출산고령사회정책 관련 전문가 조사)</td><td>2015</td></tr> <tr><td>51</td><td>2014년 저출산고령사회정책 성과평가 (저출산고령사회정책의 성과에 대한 국민 체감도 조사)</td><td>2015</td></tr> <tr><td>50</td><td>출산율 반등세로의 전환을 위한 대응전략 마련 연구</td><td>2015</td></tr> <tr><td>49</td><td>산후조리원 걱정 운영기준 마련에 관한 연구 (산후조리원 걱정운영 방안에 관한 전문가 조사)</td><td>2015</td></tr> <tr><td>48</td><td>산후조리원 걱정 운영기준 마련에 관한 연구 (산후조리원 실용면접조사)</td><td>2015</td></tr> <tr><td>47</td><td>산후조리원 걱정 운영기준 마련에 관한 연구 (산후조리원 실태조사)</td><td>2015</td></tr> <tr><td>46</td><td>2015년 전국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(미혼))</td><td>2015</td></tr> <tr><td>45</td><td>2015년 전국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(기혼))</td><td>2015</td></tr> <tr><td>44</td><td>2014년도 난임부부 지원사업 결과분석 및 평가 (난임부부 지원사업에 대한 만족도 및 요구도 조사)</td><td>2015</td></tr> </tbody> </table> << < 1 2 3 4 5 6 > >> 개인정보처리방침 이용약관 RSS 국내통계사이트 해외통계사이트 303-47 세종특별자치시 시청대로 370 세종국책연구단지 사회정책동3층(5층) TEL: 044-287-8007 FAX: 044-287-8069 E-MAIL: micro@khasa.go.kr Copyright © by KHASA. All rights Reserved. 본 홈페이지에 게시된 이메일주소로 자동으로 수집하는 것을 거부하며, 이를 위반시 관련법에 따라 처벌됨을 유념하시기 바랍니다.				번호	연구과제명(데이터명)	발행년도	53	고령화 및 생산가능인구 감소에 따른 대응전략 마련 연구	2015	52	2014년 저출산고령사회정책 성과평가 (저출산고령사회정책 관련 전문가 조사)	2015	51	2014년 저출산고령사회정책 성과평가 (저출산고령사회정책의 성과에 대한 국민 체감도 조사)	2015	50	출산율 반등세로의 전환을 위한 대응전략 마련 연구	2015	49	산후조리원 걱정 운영기준 마련에 관한 연구 (산후조리원 걱정운영 방안에 관한 전문가 조사)	2015	48	산후조리원 걱정 운영기준 마련에 관한 연구 (산후조리원 실용면접조사)	2015	47	산후조리원 걱정 운영기준 마련에 관한 연구 (산후조리원 실태조사)	2015	46	2015년 전국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(미혼))	2015	45	2015년 전국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(기혼))	2015	44	2014년도 난임부부 지원사업 결과분석 및 평가 (난임부부 지원사업에 대한 만족도 및 요구도 조사)	2015
번호	연구과제명(데이터명)	발행년도																																		
53	고령화 및 생산가능인구 감소에 따른 대응전략 마련 연구	2015																																		
52	2014년 저출산고령사회정책 성과평가 (저출산고령사회정책 관련 전문가 조사)	2015																																		
51	2014년 저출산고령사회정책 성과평가 (저출산고령사회정책의 성과에 대한 국민 체감도 조사)	2015																																		
50	출산율 반등세로의 전환을 위한 대응전략 마련 연구	2015																																		
49	산후조리원 걱정 운영기준 마련에 관한 연구 (산후조리원 걱정운영 방안에 관한 전문가 조사)	2015																																		
48	산후조리원 걱정 운영기준 마련에 관한 연구 (산후조리원 실용면접조사)	2015																																		
47	산후조리원 걱정 운영기준 마련에 관한 연구 (산후조리원 실태조사)	2015																																		
46	2015년 전국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(미혼))	2015																																		
45	2015년 전국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(기혼))	2015																																		
44	2014년도 난임부부 지원사업 결과분석 및 평가 (난임부부 지원사업에 대한 만족도 및 요구도 조사)	2015																																		

■ 화면 구성

- ‘주제별 인기자료’는 선택한 연구 주제의 조회 수가 많은 5개 자료 중 3개를 무작위로 노출함.

■ 수행 로직

- 조회 수가 많은 자료 5개 중에서 3개를 무작위로 출력하여, 같은 콘텐츠가 계속해서 노출되는 현상을 예방함.
- 사이트 운영상 필요한 경우, 관리자가 직접 인기자료 콘텐츠를 지정할 수 있으며, 관리자가 직접 지정한 인기 콘텐츠는 메인 화면 인기 콘텐츠와 같은 로직을 수행하여 70% 확률로 노출됨.
- 각 주제별로 5개 이상의 연구과제가 존재해야 로직이 생성됨.

화면명	연구DB - 주제별 - 보고서 조회	구분	수정(개선 전)
------------	----------------------------	-----------	-----------------

로그인 마이페이지 회원가입 사이트맵 사이트맵 검색어를 입력해주세요



보건복지데이터포털
Health and Welfare Data Portal

연구DB **Data Bank** **사회보장정보** **공간정보통계** **해외자료** **연로그래픽스** **정보마당**

연구 DB

주제별

- 연구/가주
 - 연구변화
 - ① **질환관리**
 - ② 고령화
 - ③ 돌싱/사망
 - ④ 결혼이혼
 - ⑤ 이민/입양
 - 성애주기
 - 가족
- 보건의료
- 사회복지
- 지역사회
 - ① 사회통합정책/심의실
 - ② 소독/재난/소비
 - ③ 빈곤/불평등
 - ④ 조사/통계/정보
 - ⑤ 종말/보존
 - ⑥ 복지국가비교
 - ⑦ 미래사회

상애주기별

영향별

주제별

🏠 연구 DB > 주제별

연구과제명	가용기 여성 임신 전 출산간접 관리지원 방안 연구(출생보조서)
발행년도	2013
연구분류코드	저출산, 여성, 건강증진
연구책임자	이상임
공동연구자	
개요	임신 전 출산간접 관리 실태 분석(한국보건사회연구원) -임신 전 출산간접 관리의 실태와 문제점을 파악하고 임신 전 출산간접 관리 증진을 위한 정책 방향을 제시하고자 함 -임신 전 출산간접 관리를 임신 위험 지시 수준과 임신 준비행동으로 구분하여 현황을 분석하고, 도출되는 문제점들의 발생 과정과 원인을 규명 고태임산부 출산간접 증진 방안 분석(한국여성정책연구원) 출산결과 이상의 고태임산부 고위험임산부를 대상으로 이들의 임신 전 출산 간접증을 분석하고, 고태 임산부에 대한 정부지급 대상화의 적절성 및 대응정책 방안 등을 개관 -5세 이상 고태 임산부의 출산결과와 질병관리에서의 위험성 여부 및 독이성을 분석하고, 위험 및 고태 임산부의 기존 및 정책대응으로서의 적절성 등을 분석
데이터 생산여부	미생산
원본링크	https://www.kihisa.net/kri/html/jsp/publication/research_view.jsp?jgno=172&eno=1664&key=writer&query=&instId&cgcat_valst=2014&content_type=1&queryStrng=cnfr%5cm%5dw%5m%A9VCZCZk9dJpdGVymF1AZ5PeydOyOpemurA==

개인정보처리방침 | 이용약관 | RSS

30347 세종특별자치시 시청대로 370 세종국책연구단지 사회정책동4층~5층) | TEL : 044-287-8007 / FAX : 044-287-0069 / E-MAIL : micro@kihisa.go.kr
 Copyright © by KIHSA All rights Reserved.
 본 홈페이지에 게시된 미래일수를 자동으로 수집하는 것을 거부하며, 이를 위반시 관련법규에 따라 처벌함을 유념하시기 바랍니다.

국내통계사이트 > 해외통계사이트

■ 화면 구성

- 게시판 형태에서 연구과제명을 클릭하면, 보고서 조회 페이지로 이동함.
- 보고서 조회 창은 ‘연구과제명’, ‘발행연도’, ‘연구분류코드’, ‘연구책임자’, ‘공동연구자’, ‘개요’, ‘데이터 생산 여부’, ‘원본링크’ 등 총 8개로 구성됨.
- ‘데이터 생산 여부’가 미생산일 경우 ‘데이터 유형’ 데이터명이 보이지 않음.

화면명	연구DB - 주제별 - 보고서 조회 (동일 저자의 다른 자료)	구분	수정(개선 후)																
 보건복지데이터포털 Health and Welfare Data Portal 연구DB Data Rank 사회보장통계 공간정보통계 해외자료 인포그래픽스 정보마당																			
연구 DB 주제별		주제별 · 연구 DB · 주제별																	
연구/가구 - 인구변화 - 인구총량 - 인구구조 - 출생/사망 - 이민/이출 - 이탈/이입 - 생애주기 - 가족 - 노년노령 - 사망/사후 - 지역사회 - 사회복지/복지/실업률 - 소득/자산/소비 - 복지/보장/통계 - 조사/통계/정보 - 통계/보통 - 복지국가비교 - 미래사회 생애주기별 명칭별		<table border="1"> <tr> <td>연구과제명</td> <td>우리나라의 AAI(Active Ageing Index) 개발 및 산출방안 기초연구</td> </tr> <tr> <td>발행년도</td> <td>2015</td> </tr> <tr> <td>연구분류코드</td> <td>고령화, 노인</td> </tr> <tr> <td>연구책임자</td> <td>최형수</td> </tr> <tr> <td>공동연구자</td> <td>오미애, 진재원, 천미경</td> </tr> <tr> <td>개요</td> <td>본 연구에서 노인 일자리 사업 및 사회활동 참여 지원 등 다양한 사업 추진에 의한 활기찬 고령사회의 변화 를 측정할 수 있는 대안적 통합 지수로서, '활기찬 고령화 통합지수 KAAI(Korean Active Ageing Index)' 개발을 위한 기초 연구 및 산출 기반 구축을 목적으로 설정함.</td> </tr> <tr> <td>데이터 생산여부</td> <td>미생산</td> </tr> <tr> <td>원본링크</td> <td>https://wish.welfare.seoul.kr/upload/data/20160425222722709.pdf</td> </tr> </table> 동일 저자의 다른 자료 • 2013년 한국복지패널 심층분석 보고서-신규 표본가구 통합 DB(KOWEPS-Combined)를 중심으로- • 2013년 서울복지실용조사 기초연구 보고서 (서울복지실용조사) • 2012년 보육료 지원대상 선정기준 조정방안 연구 (R&D) 이 자료와 같이 본 자료 • 2011년 노인실태조사 (노인실태조사) • 2013년 최저생계비 계측조사 연구 (2013년 최저생계비 계측을 위한 국민생활실태조사) • 여성노인의 노후빈곤 현황 및 대응정책 개발 (여성노인의 빈곤실태에 대한 심층연구)		연구과제명	우리나라의 AAI(Active Ageing Index) 개발 및 산출방안 기초연구	발행년도	2015	연구분류코드	고령화, 노인	연구책임자	최형수	공동연구자	오미애, 진재원, 천미경	개요	본 연구에서 노인 일자리 사업 및 사회활동 참여 지원 등 다양한 사업 추진에 의한 활기찬 고령사회의 변화 를 측정할 수 있는 대안적 통합 지수로서, '활기찬 고령화 통합지수 KAAI(Korean Active Ageing Index)' 개발을 위한 기초 연구 및 산출 기반 구축을 목적으로 설정함.	데이터 생산여부	미생산	원본링크	https://wish.welfare.seoul.kr/upload/data/20160425222722709.pdf
연구과제명	우리나라의 AAI(Active Ageing Index) 개발 및 산출방안 기초연구																		
발행년도	2015																		
연구분류코드	고령화, 노인																		
연구책임자	최형수																		
공동연구자	오미애, 진재원, 천미경																		
개요	본 연구에서 노인 일자리 사업 및 사회활동 참여 지원 등 다양한 사업 추진에 의한 활기찬 고령사회의 변화 를 측정할 수 있는 대안적 통합 지수로서, '활기찬 고령화 통합지수 KAAI(Korean Active Ageing Index)' 개발을 위한 기초 연구 및 산출 기반 구축을 목적으로 설정함.																		
데이터 생산여부	미생산																		
원본링크	https://wish.welfare.seoul.kr/upload/data/20160425222722709.pdf																		
개인정보처리방침 이용약관 RSS 국내통계사이트 해외통계사이트 30347 세종특별자치시 시청대로 370 세종국책연구단지 사회정책동(5층) TEL : 044-287-9007 FAX : 044-287-9069 E-MAIL : minso@khasa.kr Copyright © by KH-HSA. All rights Reserved. 본 홈페이지에 게시된 이메일주소물 사용으로 인하여 발생하는 모든 귀책사유를 당사자가 부담합니다.																			

■ 화면 구성

- ‘동일 저자의 다른 자료’는 선택한 연구과제 연구 책임자의 다른 연구 과제 3개를 무작위로 노출하여, 해당 저자의 다른 자료들을 쉽게 열람할 수 있도록 함.

화면명	Data Bank - 주제별	구분	수정(개선 전)
-----	-----------------	----	----------

로그인 마이페이지 통합가입 서비스 소개 서비스맵 검색어를 입력해주세요

보건복지데이터포털 연구DB Data Bank 사회보장통계 공간정보통계 해외자료 연로그레픽스 정보매당

Data Bank

주제별

- 인구/가구
 - 인구연화
 - 인구통계
 - 고령화
 - 출생/사망
 - 결혼/이혼
 - 이민/입국
- 경제주거
- 가족
- 보건의료
- 사회복지
 - 사회통계/정책/실적
 - 소득/자산/소비
 - 빈곤/불평등
 - 노년/장애/정보
 - 통계보통
 - 복지국가비교
 - 국제사회

데이터유형별

Data Summary

이용안내

신청 및 처리현황

자료문의

주제별

전체 마이크로데이터(양적) 마이크로데이터(질적) 매크로데이터

데이터를

번호	데이터명(연구과제명)	보고서 발행년도
34	남일부부 지원사업에 대한 만족도 및 요구도 조사 (2014년도 남일부부 지원사업 결과분석 및 평가)	2015
33	신후조리원 실태조사 (신후조리원 학정 운영기준 마련에 관한 연구)	2015
32	신후조리원 실용현황조사 (신후조리원 학정 운영기준 마련에 관한 연구)	2015
31	신후조리원 학정운영 방안에 관한 전문가 조사 (신후조리원 학정 운영기준 마련에 관한 연구)	2015
30	저출산 고령사회 통태적 분석을 위한 지역추적조사(3차) (저출산 고령사회 통태적 분석을 위한 지역...	2015
29	저출산 고령사회정책 관련 전문가 조사 (2014년 저출산 고령사회정책 성과평가)	2015
28	저출산 고령사회정책의 성과에 대한 국민 제언도 조사 (2014년 저출산 고령사회정책 성과평가)	2015
27	한국 출산력 및 가족보건복지 실태조사(기초) (2015년 한국 출산력 및 가족보건복지실태조사)	2015
26	한국 출산력 및 가족보건복지 실태조사(미혼) (2015년 한국 출산력 및 가족보건복지실태조사)	2015
25	미혼 여성의 여성계 건강에 대한 조사 (미혼여성 임신 전 출산건강 관리지침 개발 연구 학교 및 지역...	2014

1 2 3 4 > >>

개인정보처리방침 | 이용약관 | RSS

30547 세종특별자치시 시청대로 370 세종국책연구단지 사회정책빌딩(5층) | TEL: 044-287-6007 | FAX: 044-287-6008 | E-MAIL: mico@khis.go.kr
Copyright © by KHVISA. All rights Reserved.
본 홈페이지에 게시된 이메일주소 등 자동적으로 수신되는 것을 거부하며, 이를 위반시 관련법에 의거 처벌받을 수 있습니다.

■ 화면 구성

- 주제별 트리구조는 연구DB의 주제별 분류와 동일한 내용으로 구성됨.
- 주제별 분류체계는 연구과제의 중요단어(keyword)를 고려하여 연구진 회의에서 분류작업을 수행하였으며, 연구과제에서 생산된 각각의 데이터는 하나의 범주에만 포함되지 않고 여러 주제별 트리구조에 포함됨.
- 선택한 연구 주제에 따라 연구과제 목록이 발행년도 기준으로 정렬되어 출력됨.
- 데이터 유형을 ‘마이크로데이터(양적)’, ‘마이크로데이터(질적)’, ‘매크로데이터’로 구분함.

화면명	Data Bank - 주제별 - 인기자료	구분	수정(개선 후)																																	
 연구DB Data Bank 사회보장통계 공간정보통계 해외자료 인포그래픽스 정보마당 Data Bank 주제별 - 인구/가구 - 보건의료 - 건강보험 - 의료수급 - 질서건강 - 의료통계 - 보건산업 - 보건복지 - 사회보장/입법/실무 - 소속/가/소외 - 입원/의료통계 - 조사/가/통 - 조사/통/가/통 - 통계/통/통 - 통계통계통 - 통계통계통 주제별 주제별 인기자료 - 고위험(고위험) 임신부 정의 및 정책 수요에 대한 전문가 조사 (2013...) - 한국사회의 사회 심리적 불안 실태조사 (한국사회의 사회 심리적 불안의 원인분석과 대응방안) - 전통가, 실무자와 결혼이주여성 FGI 및 심층면담조사 (다문화가...) 주제별 주제별 최신자료 - 시군구별 건강불균등지수(Health Poverty Index) 지표 선정용 위한... - 한국사회의 사회 심리적 불안 실태조사 (한국사회의 사회 심리적 불안의 원인분석과 대응방안) - 서울 K구 지역 실태조사(2014) (노인 보건사업의 현황과 개선방안) 전체 마이크로데이터(양적) 마이크로데이터(질적) 매크로데이터 검색 <table border="1"> <thead> <tr> <th>번호</th> <th>데이터명(연구과제명)</th> <th>보고서 발행년도</th> </tr> </thead> <tbody> <tr> <td>10</td> <td>시군구별 건강불균등지수(Health Poverty Index) 지표 선정용 위한 전문가 의견 조사 (한국의 건강불균...</td> <td>2015</td> </tr> <tr> <td>9</td> <td>한국사회의 사회 심리적 불안 실태조사 (한국사회의 사회 심리적 불안의 원인분석과 대응방안)</td> <td>2015</td> </tr> <tr> <td>8</td> <td>서울 K구 지역 실태조사(2014) (노인 보건사업의 현황과 개선방안)</td> <td>2014</td> </tr> <tr> <td>7</td> <td>고위험(고위험) 임신부 정의 및 정책 수요에 대한 전문가 조사 (고위험 임신부의 산전 실태 및 출산결과와...</td> <td>2013</td> </tr> <tr> <td>6</td> <td>최근 출산여성 및 전문가 심층면담 (임신 전 출산건강 관리 실태 및 지원 방안)</td> <td>2013</td> </tr> <tr> <td>5</td> <td>임신준비 및 출산에 대한 조사 (임신 전 출산건강 관리 실태 및 지원 방안)</td> <td>2013</td> </tr> <tr> <td>4</td> <td>늦어준 주인의 의료이용 형태 및 건강증진서비스 이용욕구 조사 (늦어준지역 주인의 건강증진을 위...</td> <td>2012</td> </tr> <tr> <td>3</td> <td>전통가, 실무자와 결혼이주여성 FGI 및 심층면담조사 (다문화가족 여성과 아들의 건강실태 및 건강서...</td> <td>2012</td> </tr> <tr> <td>2</td> <td>(사회 노인장애인) 사회복지관 이용자 건강증진서비스 실태조사 (건강증진서비스 전달체계 확충방안...</td> <td>2012</td> </tr> <tr> <td>1</td> <td>보건복지서비스 연계 실태조사(기관조사) (건강증진서비스 전달체계 확충방안 연구)</td> <td>2012</td> </tr> </tbody> </table> 이전 1 다음 개인정보처리방침 이용약관 RSS 30347 세종특별자치시 시청대로 370 세종국책연구단지 사회정책동3층-5층 TEL: 044-287-8007 FAX: 044-287-8089 E-MAIL: mino@khasa.go.kr Copyright © by KHASA. All rights Reserved. 본 홈페이지에 게시된 이메일주소로 자동으로 수신하는 것을 거부하며, 이를 위반 시 관련법에 따라 처벌받을 수 있습니다.				번호	데이터명(연구과제명)	보고서 발행년도	10	시군구별 건강불균등지수(Health Poverty Index) 지표 선정용 위한 전문가 의견 조사 (한국의 건강불균...	2015	9	한국사회의 사회 심리적 불안 실태조사 (한국사회의 사회 심리적 불안의 원인분석과 대응방안)	2015	8	서울 K구 지역 실태조사(2014) (노인 보건사업의 현황과 개선방안)	2014	7	고위험(고위험) 임신부 정의 및 정책 수요에 대한 전문가 조사 (고위험 임신부의 산전 실태 및 출산결과와...	2013	6	최근 출산여성 및 전문가 심층면담 (임신 전 출산건강 관리 실태 및 지원 방안)	2013	5	임신준비 및 출산에 대한 조사 (임신 전 출산건강 관리 실태 및 지원 방안)	2013	4	늦어준 주인의 의료이용 형태 및 건강증진서비스 이용욕구 조사 (늦어준지역 주인의 건강증진을 위...	2012	3	전통가, 실무자와 결혼이주여성 FGI 및 심층면담조사 (다문화가족 여성과 아들의 건강실태 및 건강서...	2012	2	(사회 노인장애인) 사회복지관 이용자 건강증진서비스 실태조사 (건강증진서비스 전달체계 확충방안...	2012	1	보건복지서비스 연계 실태조사(기관조사) (건강증진서비스 전달체계 확충방안 연구)	2012
번호	데이터명(연구과제명)	보고서 발행년도																																		
10	시군구별 건강불균등지수(Health Poverty Index) 지표 선정용 위한 전문가 의견 조사 (한국의 건강불균...	2015																																		
9	한국사회의 사회 심리적 불안 실태조사 (한국사회의 사회 심리적 불안의 원인분석과 대응방안)	2015																																		
8	서울 K구 지역 실태조사(2014) (노인 보건사업의 현황과 개선방안)	2014																																		
7	고위험(고위험) 임신부 정의 및 정책 수요에 대한 전문가 조사 (고위험 임신부의 산전 실태 및 출산결과와...	2013																																		
6	최근 출산여성 및 전문가 심층면담 (임신 전 출산건강 관리 실태 및 지원 방안)	2013																																		
5	임신준비 및 출산에 대한 조사 (임신 전 출산건강 관리 실태 및 지원 방안)	2013																																		
4	늦어준 주인의 의료이용 형태 및 건강증진서비스 이용욕구 조사 (늦어준지역 주인의 건강증진을 위...	2012																																		
3	전통가, 실무자와 결혼이주여성 FGI 및 심층면담조사 (다문화가족 여성과 아들의 건강실태 및 건강서...	2012																																		
2	(사회 노인장애인) 사회복지관 이용자 건강증진서비스 실태조사 (건강증진서비스 전달체계 확충방안...	2012																																		
1	보건복지서비스 연계 실태조사(기관조사) (건강증진서비스 전달체계 확충방안 연구)	2012																																		

■ 화면 구성

- ‘주제별 인기자료’는 선택한 연구 주제의 조회 수가 많은 5개 자료 중 3개를 무작위로 노출함.

■ 수행 로직

- 조회 수가 많은 자료 5개 중에서 3개를 무작위로 출력하여, 같은 콘텐츠가 계속해서 노출되는 현상을 예방함.
- 사이트 운영상 필요한 경우, 관리자가 직접 인기자료 콘텐츠를 지정할 수 있으며, 관리자가 직접 지정한 인기 콘텐츠는 메인 화면 인기 콘텐츠와 같은 로직을 수행하여 70% 확률로 노출됨.
- 각 주제별로 5개 이상의 데이터가 존재해야 로직이 생성됨.

화면명	Data Bank - 주제별 - 데이터 조회 (이 자료와 같이 본 자료)	구분	수정(개선 후)																												
 연구DB Data Bank 사회보장통계 공간정보통계 해외자료 인포그래픽스 정보마당 Data Bank 주제별 인구/가구 보건의료 건강보험 생활습관병 호신건강 약물사용 보건산업 보건통계 사회복지 지역사회 사회복지/정책/실적 소득/가정/연령 임금/보통 조사/가정 조사/복지/정보 복지/가정/정보 복지/가정/정보 지역사회 데이터유행별 Data Summary 이용안내 신청 및 처리현황 자료문의 주제별 🏠 · Data Bank · 주제별 <table border="1"> <tr> <td>데이터명</td> <td>시군구별 건강빈곤지수(Health Poverty Index) 지표 선정을 위한 전문가 의견 조사</td> </tr> <tr> <td>데이터유형</td> <td>마이크로데이터(질적)</td> </tr> <tr> <td>연구과제명</td> <td>한국의 건강불평등 지표와 정책과제-건강불평등 완화를 위한 전략 연구08호기</td> </tr> <tr> <td>연구분류코드</td> <td>건강증진, 빈곤/불평등</td> </tr> <tr> <td>조사목적</td> <td>우리나라의 지역단위에서 건강불평등에 영향을 미치는 사회적 건강결정요인들을 바탕으로 지역 간 건강격차를 쉽게 확인하고 건강불평등 문제에 대한 관심을 확대하며, 이 나아가 지역 단위에서 정책의제화하는데 도움을 줄 수 있는 지표를 개발해 보고자 함.</td> </tr> <tr> <td>조사별위</td> <td></td> </tr> <tr> <td>조사방법</td> <td>E-mail 설문조사(웹파라조사)</td> </tr> <tr> <td>조사대상</td> <td>현재 건강결정요인 분야에서 연구 활동이 활발하거나 국내 건강빈곤지수 연구(신영현 등, 2009)의 참여 경험에 있는 연구자 5인</td> </tr> <tr> <td>조사기간(주기)</td> <td>2015.8.12 ~ 2015.09.14</td> </tr> <tr> <td>표본설계</td> <td>의도적 표본 추출</td> </tr> <tr> <td>통계승인여부</td> <td>미승인</td> </tr> <tr> <td>이용시유의사항</td> <td></td> </tr> <tr> <td>데이터제공여부</td> <td>불가능</td> </tr> <tr> <td>데이터관련자료</td> <td></td> </tr> </table> 이 자료와 같이 본 자료 • 2013년 한국복지패널 기초분석 보고서 (한국복지패널조사) • 빈곤에 대한 대한 생활근육구별주요 고려한 다차원성에 대한 분석 • 한국의 건강불평등 지표와 정책과제-건강불평등 완화를 위한 전략 (시군구별 건강빈곤지수(Health Poverty Index) 지표 선정을 위한 전문... 개인정보처리방침 이용약관 RSS 국내통계사이트 해외통계사이트 30347 세종특별자치시 시청대로 370 세종국립연구단지 사회정책동(5층) TEL: 044-287-8007 FAX: 044-287-8009 E-MAIL: micro@khsa.or.kr Copyright © by KHSASA. All rights Reserved. 본 홈페이지에 게시된 이메일주소로 자유롭게 수신하는 것을 거부하며, 이를 위반시 관련법에 의거 처벌됨을 유념하시기 바랍니다.				데이터명	시군구별 건강빈곤지수(Health Poverty Index) 지표 선정을 위한 전문가 의견 조사	데이터유형	마이크로데이터(질적)	연구과제명	한국의 건강불평등 지표와 정책과제-건강불평등 완화를 위한 전략 연구08호기	연구분류코드	건강증진, 빈곤/불평등	조사목적	우리나라의 지역단위에서 건강불평등에 영향을 미치는 사회적 건강결정요인들을 바탕으로 지역 간 건강격차를 쉽게 확인하고 건강불평등 문제에 대한 관심을 확대하며, 이 나아가 지역 단위에서 정책의제화하는데 도움을 줄 수 있는 지표를 개발해 보고자 함.	조사별위		조사방법	E-mail 설문조사(웹파라조사)	조사대상	현재 건강결정요인 분야에서 연구 활동이 활발하거나 국내 건강빈곤지수 연구(신영현 등, 2009)의 참여 경험에 있는 연구자 5인	조사기간(주기)	2015.8.12 ~ 2015.09.14	표본설계	의도적 표본 추출	통계승인여부	미승인	이용시유의사항		데이터제공여부	불가능	데이터관련자료	
데이터명	시군구별 건강빈곤지수(Health Poverty Index) 지표 선정을 위한 전문가 의견 조사																														
데이터유형	마이크로데이터(질적)																														
연구과제명	한국의 건강불평등 지표와 정책과제-건강불평등 완화를 위한 전략 연구08호기																														
연구분류코드	건강증진, 빈곤/불평등																														
조사목적	우리나라의 지역단위에서 건강불평등에 영향을 미치는 사회적 건강결정요인들을 바탕으로 지역 간 건강격차를 쉽게 확인하고 건강불평등 문제에 대한 관심을 확대하며, 이 나아가 지역 단위에서 정책의제화하는데 도움을 줄 수 있는 지표를 개발해 보고자 함.																														
조사별위																															
조사방법	E-mail 설문조사(웹파라조사)																														
조사대상	현재 건강결정요인 분야에서 연구 활동이 활발하거나 국내 건강빈곤지수 연구(신영현 등, 2009)의 참여 경험에 있는 연구자 5인																														
조사기간(주기)	2015.8.12 ~ 2015.09.14																														
표본설계	의도적 표본 추출																														
통계승인여부	미승인																														
이용시유의사항																															
데이터제공여부	불가능																														
데이터관련자료																															

■ 화면 구성

- ‘이 자료와 같이 본 자료’는 기존 사용자들의 과거 연구과제 열람 패턴을 분석하여, 신규 이용자에게 함께 본 자료들을 추천하는 기능임.
- 최근 3개월간 조회한 기록이 없을 경우 해당 영역을 숨김.

화면명

Data Bank - Data Summary

구분

수정(개선 후)


 보건복지데이터포털
 Health and Welfare Data Portal

연구DB

Data Bank

사회보장통계

공간정보통계

해외자료

인포그래픽스

정보마당

Data Bank

주제별

데이터유형별

Data Summary

이동안내

신청 및 처리현황

자료문의

Data Summary

☰ Data Bank - Data Summary

최신 데이터 정보

업데이트 일자 : 2016-09-27

	데이터셋	개체수 ¹⁾	변수 수	출판 연구	데이터셋 변환 날짜	현재 제공되는 데이터 조사연도
· 장애인실태조사(2014)	1	6,824	986		2016.05.02	2000 2005 2011 2014
· 환자조사(2014)	3					
- 기관조사	1	5,545	215		2016.05.03	1996 1999 2002 2005 2008 2010 2011 2012 2013 2014
- 외래환자	1	798,752	15			
- 퇴원환자	1	617,194	26			
· 한국복지패널(2015)						
- 가구조사	1	6,914	1,082	√	2016.04.07	2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
- 가구원조사	1	13,647	233			
- 부가조사	1	471	283			
· 한국의료패널(2013)						
- 가구조사	2	5,200	78	√	2016.07.12	2008 2009 2010 2011 2012 2013
- 성인가구원조사	2	14,839	99			
· 노인실태조사(2014)	1	10,451	720		2015.08.31	1994 1998 2004 2008 2011 2014
· 전국 출산력 및 가족보건복지실태조사(2015)						
- 15~49세 기혼여성조사	1	11,009	1,097		2016.07.25	1991 1994 1997 2000 2003 2006 2009 2012 2015
- 20~44세 미혼남녀조사	1	2,383	435			

1) 개체수는 가중치 반영 안한 숫자임

☰ Data Summary 목록으로

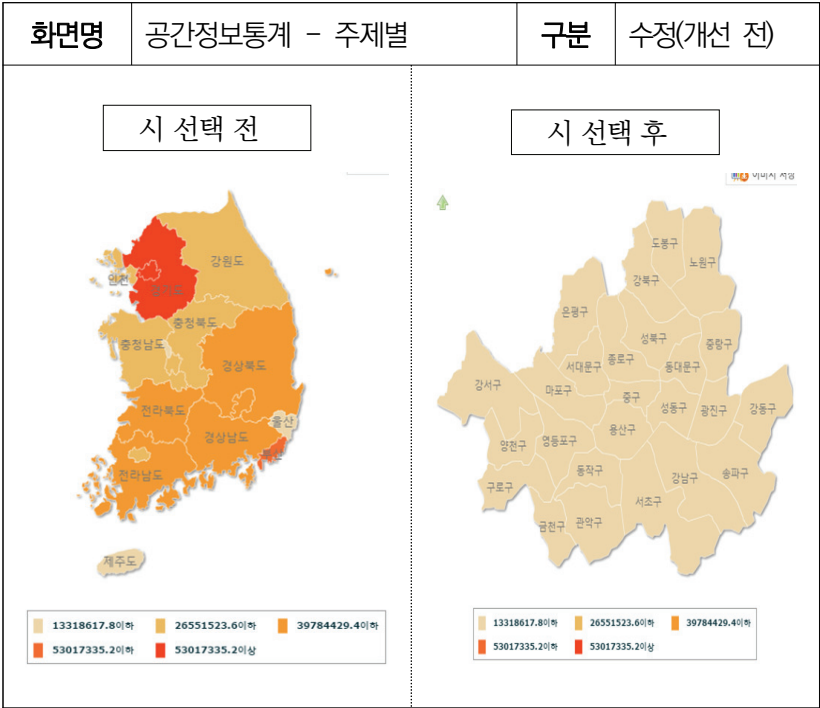
■ 화면 구성

- 한국보건사회연구원에서 수행한 대표적인 양적조사들을 한눈에 비교할 수 있는 요약 정보를 보여줌.
- 대표하는 양적조사로는 ‘장애인실태조사’, ‘환자조사’, ‘한국복지패널’, ‘한국의료패널’, ‘노인실태조사’, ‘전국 출산력 및 가족보건·복지실태조사’가 있음.
- ‘노인실태조사’를 제외한 모든 조사의 마이크로데이터가 새로 공개되었음(장애인실태조사는 2015년에 공개되었으나, 변수 추가로 인해 재공개됨).

화면명	Data Bank - Deposit Data	구분	신규 추가

■ 화면 구성

- 지금까지 보건복지데이터포털은 원내 연구보고서와 연구보고서에서 생산된 마이크로데이터를 관라·제공하고 있으나, 앞으로 원외 연구자들에게 ‘데이터 위탁관리 및 공유 서비스’를 제공하고자 함.
- ‘데이터 위탁관리 및 공유서비스’의 개요 설명, ‘데이터 위탁관리 및 공유목록’, ‘데이터 위탁관리 및 공유에 따른 혜택’, ‘데이터 위탁관리 방법 및 신청절차 안내’, ‘데이터 위탁관리 연구자의 권리 및 이익 존중’, ‘위탁관리 데이터에 대한 개인정보보호’로 구성됨.



■ 시군구 맵차트

- 시군구 통계는 맵차트 기능 미제공으로 하단의 수치에 대한 범주가 최초 계산 값 상태가 유지되어, 일개 시를 클릭하여 시군구 수치를 확인해도, 하단의 수치에 대한 범주가 시도의 범주와 같게 표시됨.
- 각 시의 데이터와 군·구 데이터의 차이가 크면 군·구 차트 색상이 모두 최하위 수치 범주의 색으로 출력되는 현상이 발생함.

화면명

공간정보통계 - 주제별

시군구 분할보기

구분

수정(개선 후)

보건복지데이터포털

Health and Welfare Data Portal

연구DB
Data Bank
사회보장통계
공간정보통계
해외자료
인포그래픽스
정보마당

공간정보통계

주제별

주제별

시군구별 현재 흡연률(현재): 2008-2014년

연도

2014

시군구 현황 보기
시군구 전체 보기

※ 각시도를 클릭하시면 해당시도의 자세한 통계정보를 보실 수 있습니다

이미지 저장

이미지 저장

22.2여

23.1여

23.9여

24.8여

20.2여

22.1여

23.9여

25.8여

통계표

시군구별 현재 흡연률(현재): 2008-2014년⁽¹⁾

단위 : %

지역	2014	
- 서울 -		21.4
충청구		24.3

■ 시군구 분할보기 기능

- 이전에 시도 지도가 보이고 일개 시를 선택해야 시군구의 지도를 확인할 수 있었던 방식에서 시도 지도와 시군구의 지도를 한 화면에 볼 수 있는 방식으로 바꾸어 구현함.

■ 범주 수처리 오류 개선

- 시군구 지도의 시도 지도의 범주 수치가 그대로 적용된 오류사항을 수정하여 시군구 지도 분할보기에서도 각각 범주 수치가 갱신되도록 구현함.
- 시를 선택할 때마다 동적으로 군·구 차트의 지도와 수치 범주를 갱신하여 각 시 안에서의 수치를 기준으로 하여 범주와 색상을 변경하여 본래 시군구 차트의 목적을 달성함.

[illegible]

■ 수행 로직

- 화면을 불러올 때, 각 시도 데이터들의 묶음과 각 시도의 시군구 데이터 묶음을 생성함.
- 좌측의 시도 차트에는 시도 데이터 묶음을 통해 데이터 범주를 계산하고 색상을 입힘.
- 사용자가 시도를 선택할 경우, 선택한 시도의 군·구 데이터 묶음을 가져와 데이터 범주를 계산하고 색상을 적용한 뒤, 새로운 맵차트를 화면 우측에 불러옴.

[illegible]

■ 기능 추가

- ‘시군구 전체보기’ 기능은 17개 시도를 각각 선택하여 상세화면에
서 시군구의 현황을 파악해야 하는 불편함을 해소하기 위하여, 한 번
에 모든 시군구의 데이터와 지도를 볼 수 있는 기능을 추가함.

화면명	웹진	구분	수정(개선 전)
보건복지통계 웹진 Health and Welfare Statistics Webzine 통계기사 알바가격 변화에 따른 인식 및 행동 변화 공연할 의사가 있는 알바 가격은 약 3,000원으로 추정되었는데, 소독수준이 높을수록 공연의사 알바가격 수 준도 높아지고 있다. 알바가격을 인상할 경우에도 고소득층보다 저소득층이 더욱 가격인상에 민감하게 반응할 것 임을 예상할 수 있어 저소득층에서 상대적으로 높은 공연 실천율을 기대할 수 있다. 공연자료제에 대한 비용보담 에 대해 국가에서 약 70%를, 본인이 30%정도 부담을 부담하는 의사가... 출처: 한국보건사회연구원, 보건복지 Issue & Focus 제21호 보건복지통계뉴스 - 작년 태어난 69세 위, 4년 살 수 있다 - 노후에 가장 중요한 것? 한국은 돈, 미국은 건강 - 남성 성인 흡연자 절반 이상, 공연의향 1 실천 1 - 신혼부부, 알바 할 · 불행 부의식적 직감 최다검색통계 - 남녀 평균수명: 1970~2009 - 시도별 위암 진료실 현황: 2011 - 부문별 복지재정규모: 2006~2012 - 사회복지예산 비중의 시·도별 현황: 2010 해외통계 - OECD, 한눈에 보는 정부지표, 2011 - OECD, Advancing Indonesian Local E-government - OECD, 중국의 경기부양책과 공공거버넌스 - OECD, OECD 국가의 공무원 인사제도 신규통계 - 북한의 GDP 추이 - 의료기관통발 의료장비 현황: 2005~2012			

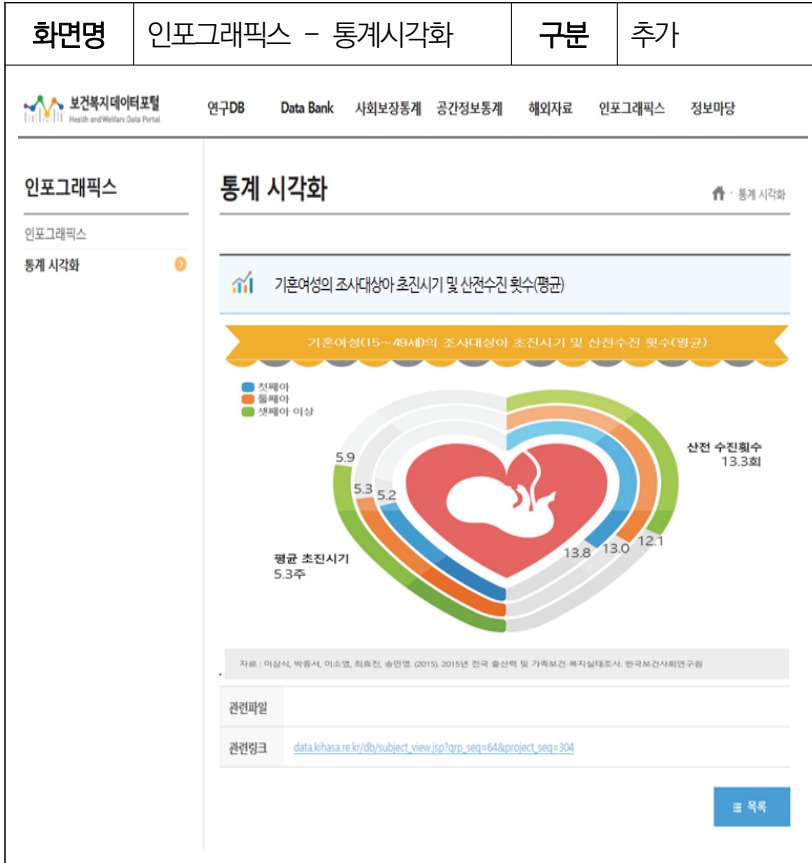
■ 화면 구성

- 정해진 콘텐츠 영역에서 정해진 글자 수 이내로 내용을 작성함.
- 스타일 지정이 불가능한 단순 텍스트로만 제작됨.
- 콘텐츠는 ‘통계기사’, ‘보건복지통계뉴스’, ‘최다검색통계’, ‘해외통계’, ‘신규통계’로 구성됨.

화면명	웹진 템플릿	구분	수정(개선 후)
	 2016.08 WEBZINE template with green header and red footer. The layout includes a title bar, a main content area with a large image, a sidebar with a search bar, and a footer with contact information.		 2016.08 WEBZINE template with orange header and blue footer. The layout includes a title bar, a main content area with a large image, a sidebar with a search bar, and a footer with contact information.
	 2016.08 WEBZINE template with red header and blue footer. The layout includes a title bar, a main content area with a large image, a sidebar with a search bar, and a footer with contact information.		 2016.08 WEBZINE template with blue header and orange footer. The layout includes a title bar, a main content area with a large image, a sidebar with a search bar, and a footer with contact information.

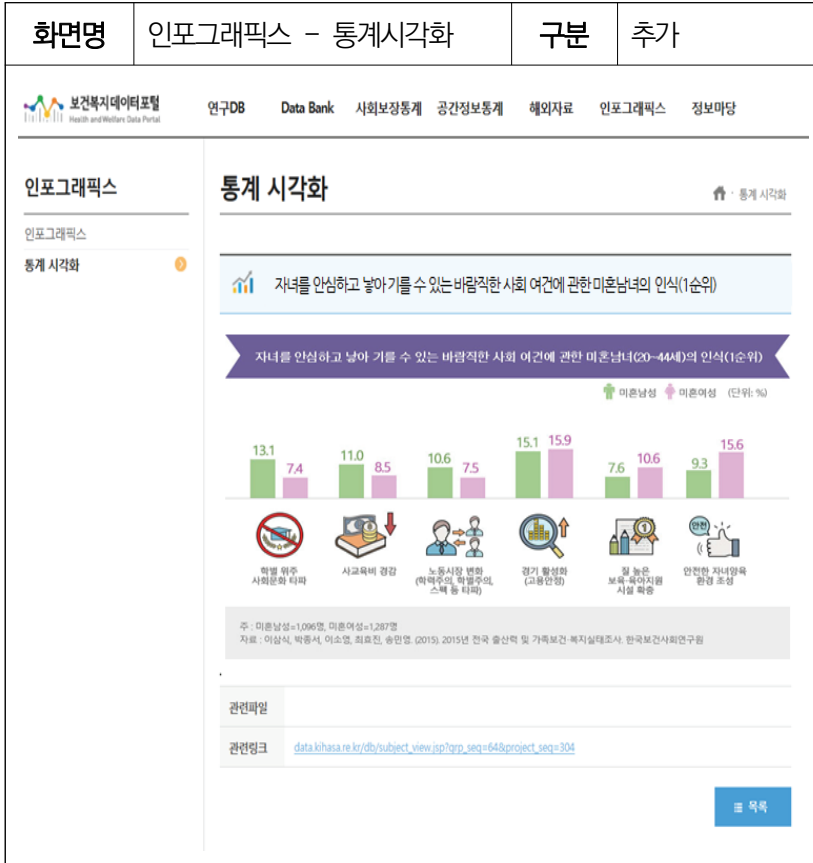
■ 웹진 헤드라인

○ 다양한 색상의 템플릿 선택 기능



통계시각화 예시

- 한국보건사회연구원 대표 조사인 ‘2015년 전국 출산력 및 가족보건·복지 실태조사-기혼여성’의 일부 결과임.
- 약 80여 건의 통계시각화 그림과 썸네일, 메인 화면 공개 디자인으로 제작되었음.



■ 통계시각화 예시

- 한국보건사회연구원 대표 조사인 ‘2015년 전국 출산력 및 가족보건·복지 실태조사-미혼남녀’의 일부 결과임.
- 약 80여 건의 통계시각화 그림과 썸네일, 메인 화면 공개 디자인으로 제작되었음.

화면명	통합 검색 결과	구분	수정(개선 전)
 통합검색 노인 Q 결과내 저장 검색도움말			
통합검색	통합검색 통합검색(317) · 연구DB(98) · Databank(4) · 사회보장통계(9) · 공간정보통계(4) 연구DB (98) 더보기 > 현재대 노인의 빈곤 실태 및 소득보장방안 연구 [2014.11.17] 이해 본 연구는 현재대 노인의 빈곤문제와 그로 인한 사회경제적 결과를 면밀히 분석하고, 이에 기반을 두어 좀 더 현실성과 실현성을 갖춘 노후 소득보장체계의 대안을 모색하는데 그 목적이 있음 📄 연구 DB > 성애주거별 > 노인 현재대 노인의 빈곤 실태 및 소득보장방안 연구 [2014.11.17] 이해 본 연구는 현재대 노인의 빈곤문제와 그로 인한 사회경제적 결과를 면밀히 분석하고, 이에 기반을 두어 좀 더 현실성과 실현성을 갖춘 노후 소득보장체계의 대안을 모색하는데 그 목적이 있음 📄 연구 DB > 주제별 > 소득/자산/소비 현재대 노인의 빈곤 실태 및 소득보장방안 연구 [2014.11.17] 이해 본 연구는 현재대 노인의 빈곤문제와 그로 인한 사회경제적 결과를 면밀히 분석하고, 이에 기반을 두어 좀 더 현실성과 실현성을 갖춘 노후 소득보장체계의 대안을 모색하는데 그 목적이 있음 📄 연구 DB > 주제별 > 빈곤/생활물 노인돌봄커리어서비스의 제공주체간 역할차별과 연계체계 구축(노인돌봄서비스 수급자 조사) [2016.03.16] 현행 노인돌봄커리어서비스를 재검토하고, 돌봄서비스의 제공주체에 대한 역할과 제공주체간 상호 연계방안을 제시하는데 목적을 두고 진행되었음. 이해 따라 공식적 돌봄서비스제도와 비공식 차등간 관계의 이론적 정립과 실적을 검토하고, 유럽 선진국가들의 돌봄서비스.. 📄 연구 DB > 성애주거별 > 노인 노인장기요양서비스의 질 관리체계 개선방안(장기요양기관대상의 설문조사) [2014.11.10] 본 연구의 목적은 장기요양서비스의 질(quality)을 개선, 향상시키기 위해 필요한 서비스 질 평가관리에 대한 전반적인 체계를 검토하고, 향후 우리 나라가 추진하여야 할 서비스의 질 평가관리체계를 개선하는데 필요한 정책적 제안을 하는데 두고 있음 📄 연구 DB > 양성별		
연구DB			
Databank			
사회보장통계			
공간정보통계			
해외자료			
★ 인기 검색어	주간 일간 1 연구 2 돌봄 3 취약계층 4 전문가 5 장애인		
★ 내가 찾은 검색어	노인		

■ 화면 구성

- 사용자가 검색한 중요단어(keyword)에 대해서 대메뉴 분류 별로 결과 제공
- 결과 내 검색기능 제공

화면명	통합 검색 결과	구분	수정(개선 후)

■ 화면 구성

- 검색 엔진 최신 버전으로 업데이트함.
- 상세 검색 기능을 추가하여, 자료의 검색 기간을 전체/최근 일주일/최근 1개월로 선택할 수 있으며, 검색 정렬을 중요도순/날짜순으로 선택하여 검색할 수 있음.

[illegible]

■ 수행 로직

- 조회 수가 많은 자료 5개 중에서 3개를 무작위로 출력하여, 같은 콘텐츠가 계속해서 노출되는 현상을 예방함.
- 사이트 운영상 필요한 경우, 관리자가 직접 인기자료 콘텐츠를 지정할 수 있음.
- 관리자가 직접 지정한 인기자료는 각 게시물마다 70% 확률로 메인 화면에 노출됨. 즉 메인 화면을 10번 접속할 경우 7번 정도 관리자가 지정한 인기자료가 메인 화면에 노출됨. 70% 확률로 노출하는 이유는 노출되는 콘텐츠가 고정되는 현상을 방지하기 위한 것임.
- 만약 관리자가 2개의 추천 콘텐츠(A, B)를 지정하면, 총 3개의 콘텐츠가 들어갈 수 있는 메인 화면 공간에, 관리자 지정 콘텐츠 A - 70%, B - 70% 확률로 노출 여부를 결정한 후에 남은 1개의 콘텐츠 자리는 조회 수가 많은 자료 5개 중에서 무작위로 선정함.

74 2016년 보건복지통계정보 통합 관리 및 운영

화면명	연구DB - 주제별 - 인가자료	구분	신규 추가
------------	--------------------------	-----------	--------------

연구DB Data Bank 사회보장통계 공간정보통계 해외자료 인포그래픽스 정보마당

연구 DB

주제별

- 인구/가구
 - 총생산
 - 고령화
 - 출생/사망
 - 죽음/이혼
 - 이동/입방
- 실업/취직
- 가족
- 보건의료
- 사회복지
- 문화/스포츠
- 환경/에너지
- 경제/산업
- 과학기술
- 외교/국제관계
- 지역개발
- 도시계획
- 교통/도로
- 농업/농촌
- 수자원/하천
- 해양수산
- 기타

생애주기별

영향범위

주제별

주제별 인가자료

- 2012년 한국 결혼 및 출산 통계조사(2012년 전국 출산력 및...)
- 가임기 여성 임신 전 출산전만 관리자를 받은 영유아발달보고서
- 임신 전 출산선과 관련 실태 및 위험요인(임산부가 및 출산자..)

주제별 최신자료

- 고령화 및 생산가능연구 감소에 따른 대응전략 마련 연구
- 2014년 저출산·고령사회정책 성과평가(저출산·고령사회정책..)
- 2014년 저출산·고령사회정책 성과평가(저출산·고령사회정책..)

번호	연구과제명(메이팅명)	발행년도
53	고령화 및 생산가능연구 감소에 따른 대응전략 마련 연구	2015
52	2014년 저출산·고령사회정책 성과평가(저출산·고령사회정책 관련 전문가 조사)	2015
51	2014년 저출산·고령사회정책 성과평가(저출산·고령사회정책의 성과에 대한 국민 제언도 조사)	2015
50	출산율 반등세로의 전환을 위한 대응전략 마련 연구	2015
49	산후조리원 적정 운영기준 마련에 관한 연구(산후조리원 적정운영 방안)에 관한 전문가 조사)	2015
48	산후조리원 적정 운영기준 마련에 관한 연구(산후조리원 실용형집주소)	2015
47	산후조리원 적정 운영기준 마련에 관한 연구(산후조리원 실태조사)	2015
46	2015년 전국 출산력 및 가족보건·복지실태조사(전국 출산력 및 가족보건·복지 실태조사(미혼))	2015
45	2015년 전국 출산력 및 가족보건·복지실태조사(전국 출산력 및 가족보건·복지 실태조사(기혼))	2015
44	2014년도 난임부부 지원사업 결과분석 및 평가(난임부부 지원사업에 대한 만족도 조사)	2015

<< < 1 2 3 4 5 6 > >>

개인정보처리방침 | 이용약관 | RSS

30247 세종특별자치시 시청대로 370 세종국립연구단지 사회정책빌딩(충·세종) | TEL : 044-387-8007 | FAX : 044-387-8069 | E-MAIL : mrc@kribs.ac.kr
Copyright © by KRISA. All Rights Reserved.

본 홈페이지에 게시된 이메일주소 등 자유로운 수집하는 것을 거부하며, 이를 위반 시 관련법령에 의거 처벌받을 수 있습니다.

■ 화면 구성

- ‘주제별 인기자료’는 선택한 연구 주제의 조회 수가 많은 5개 자료 중 3개를 무작위로 노출함.

■ 수행 로직

- 조회 수가 많은 자료 5개 중에서 3개를 무작위로 출력하여, 같은 콘텐츠가 계속해서 노출되는 현상을 예방함.
- 사이트 운영상 필요한 경우, 관리자가 직접 인기자료 콘텐츠를 지정할 수 있으며, 관리자가 직접 지정한 인기 콘텐츠는 메인 화면 인기 콘텐츠와 같은 로직을 수행하여 70% 확률로 노출됨.
- 각 주제별로 5개 이상의 연구과제가 존재해야 로직이 생성됨.

화면명	연구DB - 주제별 - 최신자료	구분	신규 추가																																	
 연구DB Data Bank 사회보장통계 공간정보통계 해외자료 인포그래픽스 정보마당																																				
연구 DB 주제별 - 인구/가부 - 인구연령 - 출생/사망 - 결혼/이혼 - 이탈/입양 - 보육/가정 - 가계 - 보건의료 - 사회복지 - 지역사회 - 사회복지/정책/실적 - 소득/자산/소비 - 의료/보장 - 주거/환경 - 조사/통계/정보 - 통일/통계 - 복지국가비교 - 미래사회	주제별 주제별 인기자료 2012년 한국 결혼 및 출산 동향 조사 (2012년 한국 출산력 및... - 가임기 여성 임신 전 출산건강 관리지원 방안 연구(통계청·보건부) - 임신 전 출산건강 관리 실태 및 지원 방안 (임신준비 및 출산에...) 주제별 최신자료 - 고령화 및 생산가능인구 감소에 따른 대응전략 마련 연구 2014년 저출산·고령사회정책 성과평가 (저출산·고령사회정책... 2014년 저출산·고령사회정책 성과평가 (저출산·고령사회정책... 연구과제명 검색 <table border="1"> <thead> <tr> <th>번호</th> <th>연구과제명(데이터명)</th> <th>발행년도</th> </tr> </thead> <tbody> <tr> <td>53</td> <td>고령화 및 생산가능인구 감소에 따른 대응전략 마련 연구</td> <td>2015</td> </tr> <tr> <td>52</td> <td>2014년 저출산·고령사회정책 성과평가 (저출산·고령사회정책 관련 전문가 조사)</td> <td>2015</td> </tr> <tr> <td>51</td> <td>2014년 저출산·고령사회정책 성과평가 (저출산·고령사회정책의 성과에 대한 국민 제갈도 조사)</td> <td>2015</td> </tr> <tr> <td>50</td> <td>출산율 변동세로의 전환을 위한 대응전략 마련 연구</td> <td>2015</td> </tr> <tr> <td>49</td> <td>산후조리원 경쟁 운영기준 마련에 관한 연구 (산후조리원 경쟁운영 발안에 관한 전문가 조사)</td> <td>2015</td> </tr> <tr> <td>48</td> <td>산후조리원 경쟁 운영기준 마련에 관한 연구 (산후조리원 실용성조사)</td> <td>2015</td> </tr> <tr> <td>47</td> <td>산후조리원 경쟁 운영기준 마련에 관한 연구 (산후조리원 실태조사)</td> <td>2015</td> </tr> <tr> <td>46</td> <td>2015년 한국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(미혼))</td> <td>2015</td> </tr> <tr> <td>45</td> <td>2015년 한국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(기혼))</td> <td>2015</td> </tr> <tr> <td>44</td> <td>2014년도 난임부부 지원사업 결과분석 및 평가 (난임부부 지원사업에 대한 만족도 및 요구도 조사)</td> <td>2015</td> </tr> </tbody> </table> 이전 1 2 3 4 5 6 다음			번호	연구과제명(데이터명)	발행년도	53	고령화 및 생산가능인구 감소에 따른 대응전략 마련 연구	2015	52	2014년 저출산·고령사회정책 성과평가 (저출산·고령사회정책 관련 전문가 조사)	2015	51	2014년 저출산·고령사회정책 성과평가 (저출산·고령사회정책의 성과에 대한 국민 제갈도 조사)	2015	50	출산율 변동세로의 전환을 위한 대응전략 마련 연구	2015	49	산후조리원 경쟁 운영기준 마련에 관한 연구 (산후조리원 경쟁운영 발안에 관한 전문가 조사)	2015	48	산후조리원 경쟁 운영기준 마련에 관한 연구 (산후조리원 실용성조사)	2015	47	산후조리원 경쟁 운영기준 마련에 관한 연구 (산후조리원 실태조사)	2015	46	2015년 한국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(미혼))	2015	45	2015년 한국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(기혼))	2015	44	2014년도 난임부부 지원사업 결과분석 및 평가 (난임부부 지원사업에 대한 만족도 및 요구도 조사)	2015
번호	연구과제명(데이터명)	발행년도																																		
53	고령화 및 생산가능인구 감소에 따른 대응전략 마련 연구	2015																																		
52	2014년 저출산·고령사회정책 성과평가 (저출산·고령사회정책 관련 전문가 조사)	2015																																		
51	2014년 저출산·고령사회정책 성과평가 (저출산·고령사회정책의 성과에 대한 국민 제갈도 조사)	2015																																		
50	출산율 변동세로의 전환을 위한 대응전략 마련 연구	2015																																		
49	산후조리원 경쟁 운영기준 마련에 관한 연구 (산후조리원 경쟁운영 발안에 관한 전문가 조사)	2015																																		
48	산후조리원 경쟁 운영기준 마련에 관한 연구 (산후조리원 실용성조사)	2015																																		
47	산후조리원 경쟁 운영기준 마련에 관한 연구 (산후조리원 실태조사)	2015																																		
46	2015년 한국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(미혼))	2015																																		
45	2015년 한국 출산력 및 가족보건복지실태조사 (한국 출산력 및 가족보건복지 실태조사(기혼))	2015																																		
44	2014년도 난임부부 지원사업 결과분석 및 평가 (난임부부 지원사업에 대한 만족도 및 요구도 조사)	2015																																		
개인정보처리방침 이용약관 RSS 국내통계사이트 해외통계사이트 30347 세종특별자치시 시청대로 370 세종국책연구단지 사회정책동(3층~5층) TEL : 044-267-8007 FAX : 044-267-8069 E-MAIL : micro@khsa.mkr Copyright © by KHS&A All rights Reserved. 본 홈페이지에 게시된 이메일주소 등 자동적으로 수집되는 것을 거부하며, 이를 위반시 관련법에 의거 처벌됨을 유념하시기 바랍니다.																																				

■ 화면 구성

- ‘주제별 최신자료’는 선택한 연구 주제의 연구과제 중 최신자료 3개를 상단에 노출하여 이용자가 최신 게시물을 쉽게 찾아갈 수 있도록 배치함.

■ 수행 로직

- 각 주제별로 5개 이상의 연구과제가 존재해야 로직이 생성됨.

화면명

연구DB - 주제별 - 보고서 조회 (동일 저자의 다른 자료)

구분

신규 추가


보건복지데이터포털
Health and Welfare Data Portal

연구DB Data Bank 사회보장통계 공간정보통계 해외자료 연로그캐릭스 정보마당

연구 DB

주제별

- [+] 연구가구
 - [+] 노년부류
 - [+] 가족상
 - [+] 노년형질
 - [+] 생애사실
 - [+] 노년이탈
 - [+] 이혼/이별
 - [+] 장애추기
 - [+] 가계
 - [+] 보건지표
 - [+] 사회통계
 - [+] 사회안전/복지/산업외
 - [+] 소액자산/소비
 - [+] 비영리활동
 - [+] 조세/지방
 - [+] 중소기업/창업
 - [+] 종교/노동
 - [+] 복지서비스요청
 - [+] 이혼사법

생애주기별

영향범

주제별

» 연구 DB - 주제별

연구과정명 우리나라의 AAI(Active Ageing Index) 개발 및 산출을 위한 기초연구

발행연도 2013

연구분류코드 고령화, 노인

연구책임자 최현수

공동연구원 오미에 진행한 현미경

개요 본 연구에서 노인 활력자 사업 및 사회통합을 위해 지원 등 다양한 사업 추진에 의해 활력이 고령사회로의 변화를 측정할 수 있는 대안적 통합 지수로서 "활기찬 고령화 통합지수 KAAI(Korean Active Ageing Index)" 개발을 위한 기초 연구 및 산출 기법 구축을 목적으로 실행함.

데이터 생산여부 대량산

원문링크 <https://wish.welfare.seoul.kr/upload/data/20160425222227209.pdf>

동일 저자의 다른 자료

- 2013년 한국복지패널 심층분석 보고서 - 신규 표본가구 통합 DB(KOWEPS_Combined)를 중심으로-
- 2013년 서울복지실태조사 기초분석 보고서 (서울복지실태조사)
- 2012년 보육률 저절대상 선결문제 해결방안 연구 (KRIH)

이 자료와 같이 본 자료

- 2011년 노인실태조사 (노인실태조사)
- 2013년 최저생계비 계측조사 연구 (2013년 최저생계비 계측을 위한 국민생활실태조사)
- 여성노인의 노후빈곤 현황 및 대응정책 개발 (여성노인의 빈곤실태에 대한 심층면접)

개발업종서명명칭 : | 이름약관 | RSS

© 2013 세종특별자치시 시청대로 370 세종국립연구단지 사회정보청(총·기획) | TEL : 044-287-8007 FAX : 044-287-8009 E-MAIL : minho@krih.ac.kr

Copyright (c) by KHASA. All rights Reserved.

본 홈페이지에 게시된 이메일주소 자로써도 수신되는 것을 거부하며, 이를 위반시 관련법령에 따라 처벌받을 수 있음을 알려드립니다.

국립보건연구원

해외보건연구원

■ 화면 구성

- ‘동일 저자의 다른 자료’는 선택한 연구 과제 연구 책임자의 다른 연구 과제 3개를 무작위로 노출하여, 해당 저자의 다른 자료들을 쉽게 열람할 수 있도록 함.

화면명

연구DB - 주제별 - 보고서 조회

(동일 저자의 다른 자료)

구분

신규 추가

국립화학정보포털
Korea Research Institute of Chemical Technology

연구DB

Data Bank

사회보장통계

공간정보통계

해외자료

인포그래픽스

장바구니

연구 DB

주제별

연구/구분

이동/복합

지출금

개발물

출발/사망

연출/이탈

입출/입장

생리/연구

기타

연구/구분

사후조사

지출금

사후조사/합계/상환금

소득/기타/소득

비용/비용

소득/비용

조사/비용/비용

비용/비용

비용/비용

비용/비용

생애주기별

연구/구분

연구/구분

연구/구분

주제별

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

연구/구분

■ 수행 로직

- 사이트 운영상 필요한 경우, 관리자가 직접 동일 저자의 다른 자료 콘텐츠를 지정할 수 있음.
- 관리자가 직접 지정한 동일 저자의 다른 자료 콘텐츠는 각 게시물마다 70% 확률로 화면에 노출됨. 즉, 메인 화면을 10번 접속할 경우 7번 정도 관리자가 지정한 콘텐츠가 화면에 노출됨. 70% 확률로 노출되는 이유는 노출되는 콘텐츠가 고정되는 현상을 방지하기 위한 것임.
- 만약 관리자가 2개의 추천 콘텐츠(A, B)를 지정하면, 총 3개의 콘텐츠가 들어갈 수 있는 메인 화면 공간에, 관리자 지정 콘텐츠 A - 70%, B - 70% 확률로 노출 여부를 결정한 후에 남은 1개의 콘텐츠 자리는 조회 수가 많은 자료 5개 중에서 무작위로 선정함.

화면명	연구DB - 주제별 - 보고서 조회 (이 자료와 같이 본 자료)	구분	신규 추가
------------	--	-----------	--------------

보건복지데이터포털
Health and Welfare Data Portal

연구DB
Data Bank
사회보장통계
공간정보통계
해외자료
인포그래픽스
정보마당

연구 DB

주제별

- 인구/가구
 - 인구변화
 - 저출산
 - 출생/사망
 - 결혼/이혼
 - 이민/입양
 - 생애주기
 - 가족
 - 보건 의료
 - 의료서비스
 - 사회복지/형식/실의료
 - 소득/재산/소비
 - 빈곤/불평등
 - 주세/자정
 - 조사/통제/경보
 - 통제/지출
 - 복지유기비교
 - 미경사회

생애주기 별

명칭 별

주제별

🏠 연구 DB · 주제별

연구과제명	우리나라의 AA(Active Ageing Index) 개발 및 산출방안 기초연구
발행년도	2015
연구분류코드	고령화, 노인
연구책임자	최현수
공동연구자	오미애, 진재원, 천미경
개요	본 연구에서 노인 일자리 사업 및 사회활동 참여 지원 등 다양한 사업 추진에 의한 활기찬 고령사회로의 변화를 측정할 수 있는 다면적 통합 지수로서, "활기찬 고령화 통합지수 KAAI(Korean Active Ageing Index)" 개발을 위한 기초 연구 및 산출 기반 구축을 목적으로 설정함.
데이터 생산여부	미생산
웹본링크	https://wish.welfare.seoul.kr/upload/data/_20160425222722709.pdf

동일 자의 다른 자료

- 2013년 한국복지패널 심층분석 보고서 - 신규 표본가구 통합 DB(KOWEPS_Combine)를 중심으로.
- 2013년 서울복지실태조사 기초분석 보고서 (서울복지실태조사)
- 2012년 부유층 지원대상 선정기준 조정방안 연구 (F&D)

이 자료와 같이 본 자료

- 2011년 노인실태조사 (노인실태조사)
- 2013년 최저생계비 계층조사 연구 (2013년 최저생계비 계층을 위한 국민생활실태조사)
- 여성노인의 노후빈곤 현황 및 대응정책 개발 (여성노인의 빈곤실태에 대한 상충면접)

개인정보처리방침 | 이용약관 | RSS

국내통계사이트 | 해외통계사이트

30347 서울특별시지치시 시청대로 370 세종국책연구단지 사립형택배(5층) | TEL: 044-287-8007 FAX: 044-287-8069 E-MAIL: micro@khasa.se.kr
Copyright © by KHASA. All rights Reserved.
본 홈페이지에 게시된 이메일주소들 자동으로 수집하는 것을 거부하며, 이를 위반시 관련법에 의거 처벌됨을 유념하시기 바랍니다.

■ 화면 구성

- ‘이 자료와 같이 본 자료’는 기존 사용자들의 과거 연구과제 열람 패턴을 분석하여, 신규 이용자에게 함께 본 자료들을 추천하는 기능임.
- 최근 3개월간 조회한 기록이 없을 경우 해당 영역을 숨김.

화면명	연구DB - 주제별 - 보고서 조회 (이 자료와 같이 본 자료)	구분	신규 추가																
 연구DB Data Bank 사회보장통계 공간정보통계 해외자료 인포그래픽스 정보마당																			
연구 DB 주제별 - 인구/가구 - 인구변화 - 저출산 - 고령/노년 - 고령/이혼 - 이민/탈탈 - 생애주기 - 가족 - 보건의료 - 사회복지 - 지역사회 - 사회통합/정책/실용 - 소득/재산/소비 - 빈곤/불평등 - 조세/자랑 - 조사/통계/정보 - 통계/연구 - 복지국가/비교 - 미래사회 생애주기별 명칭별	주제별  연구 DB · 주제별 <table border="1"> <tr> <td>연구과제명</td> <td>우리나라의 AAJ(Active Ageing Index) 개발 및 산출방안 기초연구</td> </tr> <tr> <td>발행년도</td> <td>2015</td> </tr> <tr> <td>연구분류코드</td> <td>고령화, 노년</td> </tr> <tr> <td>연구책임자</td> <td>최원수</td> </tr> <tr> <td>공동연구자</td> <td>오미애, 진재원, 천미경</td> </tr> <tr> <td>개요</td> <td>본 연구에서 노인 일자리 사업 및 사회활동 참여 지원 등 다양한 사업 추진에 의한 활기찬 고령사회로의 변화를 측정할 수 있는 대안적 통합 지수로서, '활기찬 고령화 통합지수 KAAJ(Korean Active Ageing Index)' 개발을 위한 기초 연구 및 산출 기반 구축을 목적으로 설정함.</td> </tr> <tr> <td>데이터 생산여부</td> <td>미생산</td> </tr> <tr> <td>원본링크</td> <td>https://wish.welfare.seoul.kr/upload/data/_2016042522272709.pdf</td> </tr> </table> 동일 저자의 다른 자료 2013년 한국복지패널 심층분석 보고서 - 신규 표본가구 통합 DB(KOWEPS_Combine)을 중심으로. 2013년 서울복지실태조사 기초분석 보고서 (서울복지실태조사) 2012년 보육료 지원대상 선정기준 조정방안 연구 (FGR) 이 자료와 같이 본 자료 2011년 노인실태조사 (노인실태조사) 2013년 최저생계비 계층조사 연구 (2013년 최저생계비 계층을 위한 국민생활실태조사) - 역성노년의 노후빈곤 현황 및 대응정책 개발 (역성노년의 빈곤실태에 대한 심층연구)	연구과제명	우리나라의 AAJ(Active Ageing Index) 개발 및 산출방안 기초연구	발행년도	2015	연구분류코드	고령화, 노년	연구책임자	최원수	공동연구자	오미애, 진재원, 천미경	개요	본 연구에서 노인 일자리 사업 및 사회활동 참여 지원 등 다양한 사업 추진에 의한 활기찬 고령사회로의 변화를 측정할 수 있는 대안적 통합 지수로서, '활기찬 고령화 통합지수 KAAJ(Korean Active Ageing Index)' 개발을 위한 기초 연구 및 산출 기반 구축을 목적으로 설정함.	데이터 생산여부	미생산	원본링크	https://wish.welfare.seoul.kr/upload/data/_2016042522272709.pdf		
연구과제명	우리나라의 AAJ(Active Ageing Index) 개발 및 산출방안 기초연구																		
발행년도	2015																		
연구분류코드	고령화, 노년																		
연구책임자	최원수																		
공동연구자	오미애, 진재원, 천미경																		
개요	본 연구에서 노인 일자리 사업 및 사회활동 참여 지원 등 다양한 사업 추진에 의한 활기찬 고령사회로의 변화를 측정할 수 있는 대안적 통합 지수로서, '활기찬 고령화 통합지수 KAAJ(Korean Active Ageing Index)' 개발을 위한 기초 연구 및 산출 기반 구축을 목적으로 설정함.																		
데이터 생산여부	미생산																		
원본링크	https://wish.welfare.seoul.kr/upload/data/_2016042522272709.pdf																		
개인정보처리방침 이용약관 RSS 30147 서울특별시지시 시행대로 370 세종국책연구단지 시민정책동(5층) TEL : 044-287-8007 FAX : 044-287-8069 E-MAIL : mirror@khasa.kr Copyright © by KHASA All rights Reserved. 본 홈페이지에 게시된 이메일주소들 자동으로 수집하는 것을 거부하며, 이를 위반시 관련법에 의거 처벌됨을 유념하시기 바랍니다. 국내통계사이트 해외통계사이트																			

■ 수행 로직 1

- 자바 스크립트로 클라이언트 정보를 별도의 로그 서버(웹트렌즈)로 전송하여 사용자가 조회하는 화면 정보를 기록함.
- 사용자 IP 주소와, 클라이언트마다 생성한 고유 ID를 쿠키에 저장하여, 동일한 외부 IP 주소를 갖는 네트워크 환경의 사용자들도 별도로 구분하여 데이터를 기록하기 때문에 중복을 제거하고 정확도를 높임.

화면명	연구DB - 주제별 - 보고서 조회 (이 자료와 같이 본 자료)	구분	신규 추가
------------	--	-----------	--------------

보건복지데이터포털
Health and Welfare Data Portal

연구DB
Data Bank
사회보장통계
공간정보통계
해외자료
인포그래픽스
정보마당

연구 DB

주제별

- 인구/가구
 - 인구변화
 - 인구총량
 - 인구구조
 - 인구이동
 - 인구예측
 - 생애주기
 - 가족
 - 보건의료
 - 의료서비스
 - 의료제도
 - 의료서비스
 - 의료서비스/정책/제도
 - 의료서비스/현황/분석
 - 의료서비스/연구/개발
 - 의료서비스/통계/정보
 - 의료서비스/교육/인력
 - 의료서비스/국제협력

생애주기별

명칭별

주제별 연구 DB · 주제별

연구과제명	우리나라의 AAI(Active Ageing Index) 개발 및 산출방안 기초연구
발행년도	2015
연구분류코드	고령화, 노인
연구책임자	최현수
공동연구자	오미애 진재원 천미경
개요	본 연구에서 노인 일자리 사업 및 사회활동 참여 지원 등 다양한 사업 추진에 의한 활기찬 고령사회로의 변화를 측정할 수 있는 다면적 통합 지수로서, "활기찬 고령화 통합지수 KAAI(Korean Active Ageing Index)" 개발을 위한 기초 연구 및 산출 기반 구축을 목적으로 설정함.
데이터 생산여부	미생산
원본링크	https://wish.welfare.seoul.kr/upload/data/_20160425222722709.pdf

동일 자의 다른 자료

- 2013년 한국복지패널 심층분석 보고서 - 신규 표본가구 통합 DB(KOWEPS_Combine)를 중심으로-
- 2013년 서울복지실태조사 기초분석 보고서 (서울복지실태조사)
- 2012년 보육료 지원대상 선정기준 조정방안 연구 (FG)

이 자료와 같이 본 자료

- 2011년 노인실태조사 (노인실태조사)
- 2013년 최저생계비 계측조사 연구 (2013년 최저생계비 계측을 위한 국민생활실태조사)
- 여성노인의 노후빈곤 현황 및 대응정책 개발 (여성노인의 빈곤실태에 대한 상황연결)

개인정보처리방침 | 이용약관 | RSS

30347 서울특별시지시 시행대로 370 세종국책연구단지 사당영역동(층-5층) | TEL : 044-287-8007 FAX : 044-287-8069 E-MAIL : micro@khsa.go.kr

Copyright © by KHSAGA. All rights Reserved.

본 홈페이지에 게시된 이메일주소 등 자동적으로 수집하는 것을 거부하며, 이를 위반시 관련법에 의거 처벌됨을 유념하시기 바랍니다.

국내통계사이트 | 해외통계사이트

■ 수행 로직 II

- 데이터양이 많기 때문에 데이터의 일 단위 통계 생성, 데이터 이전, URL 파싱 및 데이터 정형화, 최근 3개월 데이터 기준 추천 데이터 분석의 모든 작업은 하루 단위로 새벽 시간대에 일괄 처리됨.
- 기록된 로그 데이터를 토대로 하여, 하루 단위로 사용자가 조회한 각 콘텐츠(URL)마다, 함께 보았던 콘텐츠들의 일간 조회 횟수를 합하여 이를 로그 서버 DB에 일괄 생성함.

- 외부 서버에 저장되어 있는 하루치 통계 데이터를 가져와 URL을 파싱하여 게시판 구분, 자료 종류, 주제 등의 자료 정보를 기준 게시물, 추천 게시물 단위로 정형화하여 3개월 치 데이터를 보관함.
- 페이지 로딩 속도 향상을 위하여, 일 단위로 정형화된 데이터를 최근 3개월 동안의 기준 게시물과 추천 게시물에 대한 조회 횟수를 합쳐서, 함께 본 횟수가 많은 게시물에 대한 데이터 목록을 생성하여 저장함.
- 사용자가 특정 게시물 조회 시, 위에서 저장해 둔 기준 게시물에 대한 추천 데이터 목록을 출력함.

○ ‘주제별 인기자료’는 선택한 연구 주제의 조회 수가 많은 5개 자료 중
에서 3개를 무작위로 노출함.

- 조회 수가 많은 자료 5개 중에서 3개를 무작위로 출력하여, 같은 콘텐츠가 계속해서 노출되는 현상을 예방함.
- 사이트 운영상 필요한 경우, 관리자가 직접 인기자료 콘텐츠를 지정할 수 있으며, 관리자가 직접 지정한 인기 콘텐츠는 메인 화면 인기 콘텐츠와 같은 로직을 수행하여 70% 확률로 노출됨.
- 각 주제별로 5개 이상의 연구과제가 존재해야 로직이 생성됨.

화면명

Data Bank - 주제별 - 최신 자료

구분

신규 추가

보건복지데이터포털
Health and Welfare Data Portal

연구DB
Data Bank
사회보장통계
공간정보통계
해외자료
인로그래픽스
정보마당

Data Bank

주제별

- 인구/가구
- 보건자료
- 건강통계
- 사회복지
- 질병예방
- 질병진단
- 학술논문
- 보건산업
- 보건통계
- 사회복지
- 지역사회
- 사회통계/질병/실용
- 소득/건강/소셜
- 민간/보통
- 조사/건강
- 조사/질병/보
- 통계/보통
- 복지/건강
- 민간/보통

데이터유형별

Data Summary

이웃안내

신청 및 처리현황

자료문의

주제별

주제별 인기자료

- 고위험(고환) 임신부 정의 및 정책 수요에 대한 전문가 조사...
- 한국사회의 사회실리적 불안 실태조사 (한국사회의 사회실리적...
- 전문가 실무자와 결혼이주여성 FGI 및 심층면담조사 (다문화가...

주제별 최신자료

- 시군구별 건강빈곤지수(Health Poverty Index) 지표 선정을 위한...
- 한국사회의 사회실리적 불안 실태조사 (한국사회의 사회실리적...
- 서울 K구 지역 실태조사(2014) (노인 보건의료의 현황과 개선방...

전체

마이크로데이터(영작)

마이크로데이터(일작)

매크로데이터

데이터명

🔍

번호	데이터명(연구과제명)	보고서 발행년도
10	시군구별 건강빈곤지수(Health Poverty Index) 지표 선정을 위한 전문가 의견 조사 (한국학의 건강통계...	2015
9	한국사회의 사회실리적 불안 실태조사 (한국사회의 사회실리적 불안의 원인분석과 대응방안)	2015
8	서울 K구 지역 실태조사(2014) (노인 보건의료의 현황과 개선방안)	2014
7	고위험(고환) 임신부 정의 및 정책 수요에 대한 전문가 조사 (고환 임신부의 산전 질환 및 출산결과...	2013
6	최근 출산여성 및 전문가 실무자면담 (임신 전 출산건강 관리 실태 및 지원 방안)	2013
5	임신준비 및 출산에 대한 조사 (임신 전 출산건강 관리 실태 및 지원 방안)	2013
4	높은준비도의 의료이용 형태 및 건강관리서비스 이용욕구 조사 (높은준비도의 주인의 건강관리현황을 위...	2012
3	전문가 실무자와 결혼이주여성 FGI 및 심층면담조사 (다문화가족 여성과 아동의 건강실태 및 건강서...	2012
2	(사회·노년장애인) 사회복지자 이용자 건강관리서비스 실태조사 (건강관리서비스 전달체계 확충방안...	2012
1	보건복지서비스 연계 실태조사(기관조사) (건강관리서비스 전달체계 확충방안 연구)	2012

<<
<
1
>
>>

개인정보처리방침

이용약관

RSS

국내전화번호

해외전화번호

30347 세종특별자치시 시청대로 370 세종국립연구단지 사회정책동(5층) | TEL : 044-287-8007 | FAX : 044-287-8069 | E-MAIL : mico@khsa.go.kr

Copyright © by KHASA. All rights Reserved.

본 웹사이트에 게시된 이메일주소가 자동으로 수신되는 것을 거부하며, 이를 위반시 관련법령에 의거 처벌당할 수 있음을 알려드립니다.

■ 화면 구성

- ‘주제별 최신자료’는 선택한 연구 주제의 연구과제 중 최신자료 3개를 상단에 노출하여 이용자가 최신 게시물을 쉽게 찾아갈 수 있도록 배치함.

■ 수행 로직

- 각 주제별로 5개 이상의 연구과제가 존재해야 로직이 생성됨.

화면명	Data Bank - 주제별 - 데이터 조회 (이 자료와 같이 본 자료)	구분	신규 추가
--------------	---	-------------	----------------

보건복지데이터포털
Health and Welfare Data Portal

연구DB
Data Bank
사회보장통계
공간정보통계
해외자료
언코그래픽스
정보마당

Data Bank

- 주제별
- 인구/가구
 - 보건의료
 - 건강보험
 - 생활습관병
 - 중신상향
 - 의료비부담
 - 노년층관련
- 사회의식
 - 지역사회
 - 사회통합/질병/생태질환
 - 소속(가족/성인)
 - 의료/공공서비스
 - 진료/개업
 - 조사/조사결과/정보
 - 통계/통계자료
 - 보건관리/평가
 - 미래사회

주제별

Data Bank - 주제별

데이터명	시군구별 건강빈곤지수(health Poverty Index) 지표 선정용 위한 전문가 의견 조사
데이터유형	마이क्र로데이터(질적)
연구과제명	한국의 건강불평등 지표와 정책과제 -건강불평등을 완화하기 위한 전략 [연구DB보기]
연구분류코드	건강증진, 빈곤/불평등
조사목적	우리나라의 지역단위에서 건강불평등에 영향을 미치는 사회적 건강결정요인들을 선별하고, 지역 간 건강격차를 쉽게 확인하고 건강불평등 문제에 대한 관심을 확대하며, 더 나아가 지역 단위로 정책의제화하는데 도움을 줄 수 있는 지표를 개발해 보고자 함.
조사범위	
조사방법	E-mail 설문조사(델파이조사)
조대상	현재 건강정책성숙 분야에서 연구 활동이 활발하거나 국내 건강빈곤지수 연구실현을 위해 2009년의 참여 경험이 있는 연구자 5명
조사기간(주기)	2015.8.12 ~ 2015.09.14
표본설계	의도적 표본 추출
응답승인여부	미승인
이용사용의사항	
데이터제공여부	불가능
데이터관리자료	

이 자료와 같이 본 자료

- * 2013년 한국복지패널 기초분석 보고서 (한국복지패널조사)
- * 빈도에 대한 대안 적절근육 교육법도를 고려한 다차원성에 대한 분석
- * 한국의 건강불평등 지표와 정책과제 -건강불평등을 완화하기 위한 전략 (시군구별 건강빈곤지수(health Poverty Index) 지표 선정용 위한 전문..

개인정보처리방침 | 이용약관 | RSS

30347 세종특별자치시 시청대로 370 세종국책연구단지 사회정책플랫폼-5층 | TEL : 044-287-8007 | FAX : 044-287-8069 | E-MAIL : micro@khsa.re.kr
Copyright © by KHS-A&G All rights Reserved.

본 홈페이지에 게시된 이메일주소 등 자동적으로 수집하는 것을 거부하며, 이를 위반시 정보통신망법에 의거 처벌받을 수 있습니다.

■ 화면 구성

- ‘이 자료와 같이 본 자료’는 기존 사용자들의 과거 연구과제 열람 패턴을 분석하여, 신규 이용자에게 함께 본 자료들을 추천하는 기능임.
- 최근 3개월간 조회한 기록이 없을 경우 해당 영역을 숨김.

화면명

Data Bank – 주제별 데이터 조회 (이 자료와 같이 본 자료)

구분

신규 추가



보건복지데이터포털
Health and Welfare Data Portal

연구DB
Data Bank
사회보장통계
공간정보통계
해외자료
인포그래픽스
정보마당

Data Bank

주제별

- 인구/가구
 - 보건포털
 - 건강보험
 - 건강통계
 - 건강통계
 - 생활습관
 - 질병관리
 - 건강검진
 - 건강수준
 - 보건산업
 - 의료기기
 - 의료서비스
 - 의료서비스
 - 사후돌봄/돌봄/노년돌봄
 - 소아/청소년/소아
 - 의료/건강/의료
 - 의료/건강
 - 조사/통계/정보
 - 통계/정보
 - 보건/건강/의료
 - 보건/건강/의료

데이터유형별

Data Summary

이벤트나

신청 및 처리현황

자료문의

주제별

Data Bank : 주제별

데이터명	사군구별 건강반감지수(health Poverty Index) 지표 신청을 위한 전문가 의견 조사
데이터유형	마이프로젝트(질적)
연구과정명	한국의 건강불평등 지표와 정책과제 -건강불평등 현황을 위한 전략 연구DB보기
연구분류코드	건강통계, 보건/통계
조사목적	우리나라의 지역단위에서 건강불평등을 유발하는 사회적 건강결정요인을 바탕으로 지역 간 건강격차를 실태 파악하고, 건강불평등 완화에 대한 관심을 확대하며, 각 나라가 지역 단위에서 정책의제화하는데 도움을 줄 수 있는 자료를 개발해 보고자 함
조사방법	E-mail 설문조사(웹피아조사)
조사대상	현재 건강현황성별 분야에서 연구 활동이 활발하거나, 국내 건강반감지수 연구(신원진 등, 2009)의 참여 경험에 있는 연구자 5명
조사기간(주기)	2015.8.12 ~ 2015.09.14
표본 설계	의도적 표본 추출
응답응답여부	미응답
이용사용의사항	
데이터제공여부	불가능
데이터관련자료	

이 자료와 같이 본 자료

- 2013년 한국복지패널 기초분석 보고서 (한국복지패널조사)
- 빈곤에 대한 직접적·속구적응 고려한 디지털성에 대한 분석
- 한국의 건강불평등 지표와 정책과제 -건강불평등을 현황을 위한 전략 (사군구별 건강반감지수(health Poverty Index) 지표 신청을 위한 전문가...

개인정보처리방침 | 이용약관 | RSS

고객센터이메일
 고객센터이메일

30347 서울특별시지하철 3호선(대호로 370 서울국립연구단지 사회정책동5층-504) | TEL: 044-287-8007 | FAX: 044-287-8069 | E-MAIL: micro@khdara.kr
 Copyright (c) by KH-USA. All rights Reserved.
 본 홈페이지에 게시된 이메일주소 수신을 하는 것을 거부하여, 이를 위반시 관련법 제57조 제1항을 유증하기가 바뀔수있다.

■ 수행 로직

- 자바 스크립트로 클라이언트 정보를 별도의 로그 서버(웹트렌즈)로 전송하여 사용자가 조회하는 화면 정보를 기록함.
- 사용자 IP 주소와, 클라이언트마다 생성한 고유 ID를 쿠키에 저장하여, 동일한 외부 IP 주소를 갖는 네트워크 환경의 사용자들도 별도로 구분하여 데이터를 기록하기 때문에 중복을 제거하고 정확도를 높임.

화면명

Data Bank - 주제별 - 데이터 조회 (이 자료와 같이 본 자료)

구분

신규 추가

보건복지데이터포털
Health and Welfare Data Portal

연구DB
Data Bank
사회보장통계
공간정보통계
해외자료
인포그래픽스
정보마당

Data Bank

주제별

- 인구/가구
 - 보건보험
 - 건강보험
 - 주거환경
 - 사회보장급여
 - 국민연금
 - 직장연금
 - 국민연금
 - 보건산업
 - 보건산업
- 사회복지
 - 사회복지/정책/서비스
 - 소득/자산/소비
 - 의료/보육/돌봄
 - 노년/장애
 - 조사/통계/정보
 - 문화/스포츠
 - 지역사회

데이터유형별

Data Summary

이웃안내

신청 및 처리현황

자료문의

주제별

Data Bank - 주제별

데이터명	사건구별 건강빈곤지수(Health Poverty Index) 지표 선정용 위한 전문가 의견 조사
데이터유형	마이크로데이터(집합)
연구과제명	한국의 건강불평등 지표와 정책과제 -건강불평등을 원용을 위한 연구 연구DB보기
연구분류코드	건강질환, 빈곤/불평등
조사목적	우리나라의 지역단위에서 건강불평등을 설명할 미치는 사회적 건강결정요인들을 선별하고 지역 간 건강격차를 쉽게 확인하고, 건강불평등 문제에 대한 관심을 확대하며, 더 나아가 지역 단위에서 정책의제화하는데 도움을 줄 수 있는 지표를 개발해 보고자 함.
조사범위	
조사방법	E-mail 설문조사(웹/이메일)
조사대상	현재 건강영향성평가 분야에서 연구 활동이 활발하거나, 국내 건강빈곤지수 연구(신청된, 2009년)의 참여 경험에 있는 연구자 5인
조사기간(주기)	2015.8.12 ~ 2015.09.14
표본설계	의도적 표본 추출
통계승인여부	미승인
이용시유의사항	
데이터제공여부	불가능
데이터관련자료	

이 자료와 같이 본 자료

- 2013년 한국복지패널 기초분석 보고서 (한국복지패널조사)
- 빈곤에 대한 대안적 접근은 욕구충족을 고려한 다차원성에 대한 분석
- 한국의 건강불평등 지표와 정책과제 -건강불평등을 원용을 위한 연구 (사건구별 건강빈곤지수(Health Poverty Index) 지표 선정용 위한 전문가 의견 조사)

개인정보처리방침
이용약관
RSS

30347 세종특별자치시 시청대로 370 세종국립연구단지 사회정책빌딩(5층) | TEL: 044-287-8007 | FAX: 044-287-8069 | E-MAIL: micro@khsa.re.kr
Copyright © by KHS-AdA. All rights Reserved.

본 홈페이지에 게시된 이메일주소로 자동으로 수신하는 것을 거부하며, 이를 위반시 관련법규에 의하여 처벌을 받습니다.

■ 수행 로직 II

- 데이터양이 많기 때문에 데이터의 일 단위 통계 생성, 데이터 이전, URL 파싱 및 데이터 정형화, 최근 3개월 데이터 기준 추천 데이터 분석의 모든 작업은 하루 단위로 새벽 시간대에 일괄 처리됨.
- 기록된 로그 데이터를 토대로 하여, 하루 단위로 사용자가 조회한 각 콘텐츠(URL)마다, 함께 보았던 콘텐츠들의 일간 조회 횟수를 합하여 이를 로그 서버 DB에 일괄 생성함.

- 외부 서버에 저장되어 있는 하루치 통계 데이터를 가져와 URL을 파싱하여 게시판 구분, 자료 종류, 주제 등의 자료 정보를 기준 게시물, 추천 게시물 단위로 정형화하여 3개월 치 데이터를 보관함.
- 페이지 로딩 속도 향상을 위하여, 일 단위로 정형화된 데이터를 최근 3개월 동안의 기준 게시물과 추천 게시물에 대한 조회 횟수를 합쳐서, 함께 본 횟수가 많은 게시물에 대한 데이터 목록을 생성하여 저장함.
- 사용자가 특정 게시물 조회 시, 위에서 저장해 둔 기준 게시물에 대한 추천 데이터 목록을 출력함.

2. Deposit Data

[illegible]

■ 화면 구성

- 지금까지 보건복지데이터포털은 원내 연구보고서와 연구보고서에서 생산된 마이크로데이터를 관리·제공하고 있으나, 앞으로 원외 연구자들에게 ‘데이터 위탁관리 및 공유 서비스’를 제공하고자 함.
- ‘데이터 위탁관리 및 공유서비스’의 개요 설명, ‘데이터 위탁관리 및 공유목록’, ‘데이터 위탁관리 및 공유에 따른 혜택’, ‘데이터 위탁관리 방법 및 신청절차 안내’, ‘데이터 위탁관리 연구자의 권리 및 이익 존중’, ‘위탁관리 데이터에 대한 개인정보보호’로 구성됨.

제 4 장

보건복지데이터포털 고도화 방향: 추천 시스템

제1절 Collaborative recommendation

제2절 Content-based recommendation

제3절 소결

4

보건복지데이터포털 고도화 << 방향: 추천 시스템

제1절 Collaborative recommendation

Collaborative recommendation 방법은 사용자의 아이템에 대한 과거의 선호도가 현재 혹은 미래에도 비슷한 패턴으로 유지될 것이라고 가정하고 이를 기반으로 하여 선호도 패턴이 비슷한 사용자들의 선호도 정보를 활용하여 선호도를 예측하는 방식의 추천 시스템이다(Su and Khoshgoftaar, 2009; Ricci et al., 2011).

Collaborative recommendation 방법의 특징은 추천 시스템을 구축하는 과정에서 오직 사용자와 아이템 상호 간의 관계에만 초점을 맞춘다는 점이며 사용자나 아이템이 가지는 자체적인 특징은 추천 시스템 구축 과정에서 일체 배제된다. 따라서 데이터를 구축하는 과정이 매우 쉽다는 장점을 가지지만 사용자와 아이템의 특징을 고려한 맞춤형 추천 시스템을 구축하기는 어렵다는 단점이 있다.

Collaborative recommendation 방법에 사용되는 데이터는 다양한 형태를 가질 수 있지만 독자의 편의와 정보전달의 정확성을 위해 일반적으로 데이터가 사용자(user)와 아이템 두 가지 키워드로 구분(indexing)되는 구조를 가진다고 가정한다. 보건복지데이터포털의 경우 포털에서 제공하는 다양한 데이터 관련 서비스가 일종의 아이템에 해당하며 포털을 사용하는 사용자에게 대하여 각각의 서비스에 대한 선호도를 조사하여 데이터를 구축할 수 있을 것이다.

■ Collaborative recommendation 방법의 데이터 index

$U = \{u_1, u_2, \dots, u_n\}$: n 명의 사용자 집합

$P = \{p_1, p_2, \dots, p_m\}$: m 개의 아이템 집합

1. User-based nearest neighbor recommendation

User-based nearest neighbor recommendation 방법은 가장 기본적인 Collaborative recommendation 방법이며 다음과 같이 5단계의 절차로 이루어진다.

Step 1. 데이터의 구성

$R = (r_{u,p})_{u \in U, p \in P}$: $n \times m$ 크기의 선호도(rating) 행렬,

- $r_{u,p}$ = 사용자 $u \in U$ 의 아이템 $p \in P$ 에 대한 선호도

선호도 행렬 R 은 선호도 $r_{u,p}$ 를 행 방향으로 사용자, 열 방향으로 아이템이 위치하도록 구성한 행렬이다. 일반적으로 선호도 $r_{u,p}$ 는 0 혹은 양의 실수를 값으로 가지며 사용자의 선호도가 클수록 큰 값을 가지도록 구성한다. 사용자의 선호도가 없는(missing) 경우 편의상 $r_{u,p} = \infty$ 로 정의한다.

Step 2. 사용자 $u \in U$ 와 $v \in U$ 의 선호도의 유사도(similarity) 정의

$$sim(u, v) = \frac{\sum_{p \in P} (r_{u,p} - \bar{r}_u)(r_{v,p} - \bar{r}_v)}{\sqrt{\sum_{p \in P} (r_{u,p} - \bar{r}_u)^2} \sqrt{\sum_{p \in P} (r_{v,p} - \bar{r}_v)^2}},$$

- $\bar{r}_u = \frac{1}{m} \sum_{p \in P} r_{u,p}$: 사용자 u 의 선호도의 평균

위 유사도를 서로 다른 두 사용자 u 와 v 의 Pearson's correlation coefficient 혹은 adjusted cosine similarity라고 한다. 이 유사도는 -1과 1 사이의 값을 가지며 절댓값이 1에 가까울수록 u 와 v 의 선호도 사이에 직선관계가 강하다는 것을 의미한다. 특히 이 유사도가 양수이면 u 의 선호도가 v 의 선호도와 양의 비례관계에 있다는 것을 의미하므로, v 의 선호도가 크면 u 의 선호도도 클 것이라고 예측할 수 있다.

참고로 Pearson's correlation coefficient 이외에도 단순히 Euclidean distance를 사용할 수도 있다. 선호도의 단위가 서로 다르게 조사된 경우 Spearman's correlation과 Kendall's tau와 같이 선호도의 순위를 활용한 비모수적(non-parametric) 유사도를 사용할 수도 있다(Bonett and Wright, 2000).

Step 3. 사용자 $u \in U$ 의 k -nearest neighbor 구성

$$N_k(u) \subset U,$$

- $|N_k(u)| = k$, $|A|$ 는 집합 A 의 원소의 개수
- $\text{sim}(u, v) \geq \text{sim}(u, w)$, $v \in N_k(u)$, $w \in N_k(u)^c$
- $\text{sim}(u, v) \geq 0$, $v \in N_k(u)$

사용자 u 의 k -nearest neighbor $N_k(u)$ 는 위에서 정의한 유사도를 크기순으로 나열하였을 때 큰 값으로부터 k 번째 순위까지의 유사도를 가지는 사용자 v 의 집합을 의미한다. 특히 선호도가 비슷한 사용자들의 정보를 활용하여 추천 시스템을 구축하려는 것이므로 유사도의 값이 양수인 경우만 생각하며 따라서 크기가 k 개 이하가 될 수도 있다.

유사도와 k -nearest neighbor의 정의는 User-based nearest

neighbor recommendation 방법에서 가장 핵심적인 과정이다. 특히 아이টে에 대한 선호도가 대부분 비어 있는 사용자의 경우 유사도의 정확성이 매우 낮아지며 전체 추천 시스템의 예측 성능도 동시에 낮아진다. 이러한 경우에 어떠한 유사도를 사용하는 것이 합리적인지에 대한 다양한 연구가 진행되고 있다(Herlocker et al., 2002; Herlocker et al., 2004).

Step 4. 사용자 $u \in U$ 의 아이টে $p \in P$ 에 대한 선호도 예측

$$pred(u, p) = \bar{r}_u + \frac{\sum_{v \in N_k(u)} sim(u, v)(r_{v, p} - \bar{r}_v)}{\sum_{v \in N_k(u)} sim(u, v)}$$

사용자 u 의 k -nearest neighbor $N_k(u)$ 가 결정되면 k -nearest neighbor에 속하는 사용자 v 의 아이টে p 에 대한 선호도의 편차 $r_{v, p} - \bar{r}_v$ 를 계산하고 이를 위 2)에서 정의한 유사도 $sim(u, v)$ 를 가중치로 사용하여 가중 평균을 계산한다. 이 값은 u 와 비슷한 선호도를 가지는 사용자들이 아이টে p 에 대하여 평균적으로 얼마나 선호도의 편차를 가지는가를 계산하는 것이다. 따라서 사용자 u 의 아이টে p 에 대한 선호도는 사용자 u 의 선호도의 평균과 이 편차의 가중 평균의 합으로 정의된다.

Step 5. 사용자 $u \in U$ 에 대한 추천 시스템 구축

위 1)에서 4)까지의 과정을 거치면 시스템 내의 모든 아이টে에 대하여 사용자 u 의 선호도를 예측할 수 있고, 그에 따라 선호도가 높은 순으로 아이টে를 정렬하여 정리하면 최종적인 추천 시스템을 구축할 수 있다.

선호도 행렬 R 에서 사용자 간의 유사도를 계산하고 2-nearest

neighbor를 사용하여 사용자 u_1 의 아이템 p_5 에 대한 선호도를 예측하는 예시는 다음과 같다(Jannach et al., 2011).

$$R = \begin{pmatrix} 5 & 3 & 4 & 4 & \infty \\ 3 & 1 & 2 & 3 & 3 \\ 4 & 3 & 4 & 3 & 5 \\ 3 & 3 & 1 & 5 & 4 \\ 1 & 5 & 5 & 2 & 1 \end{pmatrix} \Rightarrow \bar{r}_1 = 4, \bar{r}_2 = 2.4 \Rightarrow \text{sim}(u_1, u_2) = 0.85$$

$$\text{sim}(u_1, u_2) = \frac{(5 - \bar{r}_1)(3 - \bar{r}_2) + \dots + (4 - \bar{r}_1)(3 - \bar{r}_2)}{\sqrt{(5 - \bar{r}_1)^2 + \dots + (4 - \bar{r}_1)^2} \sqrt{(3 - \bar{r}_2)^2 + \dots + (3 - \bar{r}_2)^2}} = 0.85$$

$$\text{sim}(u_1, u_3) = 0.7, \text{sim}(u_1, u_4) = 0.0, \text{sim}(u_1, u_5) = -0.79$$

$$\text{pred}(u_1, p_5) = 4 + \frac{0.85 \times (3 - 2.4) + 0.70 \times (5 - 3.8)}{(0.85 + 0.7)} = 4.87$$

2. Item-based nearest neighbor recommendation

위에서 소개한 User-based nearest neighbor recommendation 방법은 사용자의 선호도 사이의 유사도를 기준으로 하여 추천 시스템을 결정하는 방법이다. 따라서 시스템에 기록된 사용자의 수가 수십만 혹은 수백만에 이르고 추천 시스템을 실시간으로 유지하고자 하는 경우 추천 시스템을 구축하기 위한 유사도를 계산하고 이로부터 k -nearest neighbor를 구성하는 것이 현실적으로 가능한지에 대한 기술적인 검토가 필요하다. 만약 시스템에서 사용자의 수보다 아이템의 수가 상대적으로 적은 경우라면 사용자 사이의 유사도가 아닌 아이템 사이의 유사도를 활용한 추천 시스템을 구축하는 것이 더 효율적이다. 이러한 방법을 일반적으로 Item-based nearest neighbor recommendation 방법이라고 하며, 이는 User-based nearest neighbor recommendation 방법과

기본적으로 동일한 시스템이라고 할 수 있다.

Step 1. 아이템 $p \in P$ 와 아이템 $q \in P$ 의 선호도의 유사도 정의

$$sim(p, q) = \frac{\sum_{u \in U} (r_{u,p} - \bar{r}_p)(r_{u,q} - \bar{r}_q)}{\sqrt{\sum_{u \in U} (r_{u,p} - \bar{r}_p)^2} \sqrt{\sum_{u \in U} (r_{u,q} - \bar{r}_q)^2}},$$

$$- \bar{r}_p = \frac{1}{n} \sum_{u \in U} r_{u,p}: \text{아이템 } p \text{의 선호도의 평균}$$

Step 2. 아이템 $p \in P$ 의 k -nearest neighbor 구성

$$N_k(p) \subset P,$$

$$- |N_k(p)| = k$$

$$- sim(p, q) \geq sim(p, r), q \in N_k(p), r \in N_k(p)^c$$

$$- sim(p, q) \geq 0, q \in N_k(p)$$

Step 3. 사용자 $u \in U$ 의 아이템 $p \in P$ 에 대한 선호도 예측

$$pred(u, p) = \frac{\sum_{q \in N_k(p)} sim(p, q) r_{u,q}}{\sum_{q \in N_k(p)} sim(p, q)}$$

3. Model-based recommendation

위에서 소개한 두 가지 추천 시스템은 대용량 데이터를 대상으로 하는 경우 데이터의 저장, 처리, 계산 속도와 같은 기술적인 문제가 발생하기

쉽다. Model-based recommendation 방법은 이러한 단점을 보완하기 위해 개발된 방법으로, 먼저 추천 시스템에 대한 적당한 모형을 가정하고 주어진 데이터를 바탕으로 하여 사전에 모형을 구축한 후 추천 시스템을 적용할 때는 구축된 모형을 활용하는 방식을 취한다. 따라서 새로운 모형을 구축할 필요가 있다고 판단되기 전까지 실시간으로 추천 시스템을 유지할 수 있다는 장점을 가지지만 데이터를 설명하는 모형에 대한 의존도가 높기 때문에 모형에 대한 사전 검토가 충분히 이루어진 경우에 사용하는 것이 좋다.

가. Matrix factorization(Latent factor models)

Matrix factorization 방법은 선호도 행렬을 분해하여 Latent factor를 구성하고 중요한 Latent factor로 생성되는 저차원 공간에 사용자 혹은 아이템의 선호도를 정사영하여 다른 사용자의 선호도와와의 관계를 시각적으로 표현해 주는 방법이다(Deerwester et al., 1990; Koren et al., 2009). 특히 Latent factor를 구성하는 방법은 선호도 행렬에 대한 Singular value decomposition 방법이 자주 이용된다(Sarwar et al., 2000a; Grigorik, 2007).

Matrix factorization 방법은 추천 시스템을 구성하는 과정에서 다양한 방식으로 응용될 수 있다. 예를 들면 정사영 공간에서 적당한 거리(distance)를 정의하면 특정 사용자로부터 가까운 거리에 위치하는 사용자들의 집합을 정의할 수 있고 이러한 사용자들의 집합을 k -nearest neighbor로 지정하면 User-based nearest neighbor recommendation 방법을 직접 사용할 수 있다. Matrix factorization 방법을 적용하기 위한 데이터는 선호도 행렬이며 절차는 다음과 같다.

Step 1. 선호도 행렬 R^T 의 분해

$$R^T = P \Sigma U^T$$

선호도 행렬 R 에 대하여 위 등식을 만족하는 세 개의 행렬 P , Σ , U 를 찾는 방법을 R^T 의 Singular value decomposition이라고 한다. 위 등식에서 행렬 Σ 는 크기가 $m \times n$ 이고 비대각(off-diagonal) 원소가 모두 0으로 구성된 행렬이며 Σ 의 대각원소를 R^T 의 Singular value라고 한다. 행렬 U 는 사용자에게 대응하는 Singular vector로 구성된 행렬이며 크기가 $n \times n$ 이고 행렬 P 는 아이템에 대응하는 Singular vector로 구성된 $m \times m$ 크기의 행렬이다. Matrix factorization 방법에서는 Singular vector를 Latent vector로 해석한다.

Step 2. 선호도 벡터 R_u 와 R_p 의 저차원 정사영 구성

$$R_{u,2} = R_u P_2 \Sigma_2^{-1},$$

- R_u =사용자 u 에 대응하는 선호도 행렬 R 의 행벡터
- Σ_2 =값이 가장 큰 2개의 Singular value만을 포함하는 크기 2×2 인 Σ 의 부분행렬(sub-matrix)
- P_2 = Σ_2 가 포함하는 2개의 Singular value에 대응하는 2개의 Singular vector를 포함하는 P 의 부분행렬

$$R_{p,2} = R_p^T U_2 \Sigma_2^{-1},$$

- R_p =아이템 p 에 대응하는 선호도 행렬 R 의 열벡터
- U_2 = Σ_2 가 포함하는 2개의 Singular value에 대응하는 2개의 Singular vector를 포함하는 U 의 부분행렬

위 정의에서 벡터 $R_{u,2}$ 는 벡터 R_u 를 부분행렬 P_2 의 열벡터가 생성하는 2차원 공간에 정사영한 정사영 벡터를 의미한다. 따라서 모든 사용자의 정사영 벡터를 계산하여 2차원으로 도시하면 사용자 u 와 비슷한 위치에 있는 사용자의 집합을 시각적으로 쉽게 확인할 수 있다. 물론 이러한 저차원 공간에서 적당한 거리 혹은 유사도를 정의하여 u 와 가까운 k -nearest neighbor를 수치적으로 엄밀하게 구성할 수도 있다.

선호도 행렬 R 을 분해하여 선호도 벡터 $R_u = (5, 3, 4, 4)$ 의 P_2 로 생성된 부분공간에서의 좌표를 계산하면, 벡터 $R_{u,2}$ 는 다음과 같다(Jannach et al., 2011).

$$R = \begin{pmatrix} 3 & 1 & 2 & 3 \\ 4 & 3 & 4 & 3 \\ 3 & 2 & 1 & 5 \\ 1 & 6 & 5 & 2 \end{pmatrix} \Rightarrow P_2 = \begin{pmatrix} -0.4312 & 0.4931 \\ -0.5327 & -0.5305 \\ -0.5237 & -0.4052 \\ -0.5058 & 0.5578 \end{pmatrix}, \Sigma_2 = \begin{pmatrix} 12.2215 & 0 \\ 0 & 4.9282 \end{pmatrix}$$

$$\Rightarrow R_{u,2} = R_u \times P_2 \times \Sigma_2^{-1} = (-0.64, 0.03).$$

나. Association rule mining

Association rule mining 방법은 대표적인 Data mining 방법의 하나로 사용자의 아이템구매 패턴 간의 연관관계를 구축하는 방법이다 (Sarwar et al., 2000b; Lin, 2000; Lin et al., 2002; Fu et al., 2000; Mobasher et al., 2001; Shyu et al., 2005). 예를 들어 유아용 식품을 구매한 사용자들의 70%는 기저귀를 구매한다는 연관관계는 유아용 식품과 기저귀를 대상으로 하여 다양한 마케팅 전략을 구축할 수 있도록 해준다. 하지만 시스템에 기록된 아이템의 목록이 다양할수록 연관관계 규칙은 기하급수적으로 늘어나므로 유용한 연관규칙을 효율적으로 추출할 수

있는 알고리즘이 필요하다. 참고로 아이тем 간의 Association rule이 아니라 사용자 간의 Association rule을 사용할 수도 있다.

Step 1. 데이터의 구성

$S = (s_{u,p})_{u \in U, p \in P} : n \times m$ 구매내역 행렬,

- $s_{u,p}$ = 사용자 $u \in U$ 의 아이тем $p \in P$ 에 대한 구매내역

구매내역 행렬 S 는 구매내역 $s_{u,p}$ 를 행 방향으로 사용자, 열 방향으로 아이тем이 위치하도록 구성한 행렬이며 구매내역 $s_{u,p}$ 는 사용자 u 가 아이тем p 를 구매한 경우 1, 그렇지 않은 경우 0으로 정의된다.

Association rule에 필요한 정의는 다음과 같다.

■ 아이тем집합 $X \subset P$ 에서 아이тем집합 $Y \subset P$ 로의 연관규칙

$X \Rightarrow Y$,

- $X \neq \emptyset, Y \neq \emptyset, X \cap Y = \emptyset$.

■ 아이тем집합 $X \subset P$ 의 Support

$supp(X) = \sum_{u \in U} \sum_{p \in X} I(s_{u,p} = 1)$,

- I =지시(indicator)함수: $s_{u,p} = 1$ 이면 $I(s_{u,p} = 1) = 1$ 이고, 그렇지 않으면 $I(s_{u,p} = 1) = 0$ 임.

■ 연관규칙 $X \Rightarrow Y$ 의 Confidence

$$conf(X \Rightarrow Y) = \frac{supp(X \cup Y)}{supp(X)}$$

■ 연관규칙 $X \Rightarrow Y$ 의 Lift

$$lift(X \Rightarrow Y) = \frac{conf(X \Rightarrow Y)}{supp(Y)}$$

위에서 정의한 Support, Confidence, Lift는 유용한 연관규칙을 정의하기 위한 측도(measure)로 사용된다. $supp(X)$ 는 아이тем집합 X 를 구매한 사용자의 수를 의미한다. 따라서 $supp(X)$ 값이 크다는 것은 아이тем집합 X 가 판매율이 높은 아이тем들로 구성된 아이тем집합이라는 것을 의미한다. $conf(X \Rightarrow Y)$ 는 아이тем집합 X 의 구매를 조건으로 하는 아이тем집합 Y 의 구매에 대한 조건부 확률이다. 따라서 $conf(X \Rightarrow Y)$ 값이 크면 아이тем집합 X 를 구매한 사용자가 아이тем집합 Y 를 구매할 확률이 높다는 것을 의미한다. $lift(X \Rightarrow Y)$ 는 아이тем집합 X 와 Y 의 구매 패턴 간의 독립성의 정도를 의미한다. $lift(X \Rightarrow Y)$ 값이 1에 가까우면 X 와 Y 의 구매패턴이 상호 독립적임을 의미하며 1보다 커질수록 상호 의존성이 커짐을 의미한다.

Step 2. 아이тем집합 리스트 작성

Association rule mining 방법을 적용하여 유용한 연관규칙을 찾기 위해서는 사전에 아이тем집합의 리스트를 구성해야 한다. 만약 시스템 상의 아이тем의 종류가 m 개이고 가능한 모든 아이тем집합을 이 리스트에 포

함한다면 모두 2^m 개의 아이тем집합을 포함하는 리스트를 얻을 수 있다. 따라서 이 리스트에 포함된 아이тем집합의 개수는 아이тем의 개수에 대하여 기하급수적으로 증가하므로, 포함된 모든 아이тем집합을 조사하는 것은 기술적으로 불가능하다. 결국 이 리스트에 포함된 아이тем집합 중 추천 시스템 구축에 반드시 필요한 아이тем집합을 정의하고 이러한 아이тем만을 포함하는 새로운 아이тем집합 리스트를 구성해야 한다. 아래에 기술하는 아이тем집합 리스트 작성 알고리즘은 연관규칙을 구성하는 가장 대표적인 알고리즘인 Apriori 알고리즘(Agrawal, R. and Srikant, R., 1994; Bayardo Jr, R. J., 1998)의 일부이다.

■ 아이тем집합 리스트 작성 알고리즘

- [1] 아이тем 구매횟수의 하한(Lower bound) c 를 결정한다.
- [2] 아이тем집합의 집합 $L_1 = \{X \subset P : \text{supp}(X) \geq c, |X| = 1\}$ 을 구성한다.
- [3] 아래의 [3.1]에서 [3.2]의 구성과정을 반복하다가 적당한 k 에 대하여 $L_k = \{X \subset P : \text{supp}(X) \geq c, |X| = k\} = \emptyset$ 이면 구성과정을 멈춘다.
 - [3.1] L_{k-1} 에 포함된 아이тем집합들을 Inner-join하여 L_k 를 구성한다.
 - [3.2] L_k 에 포함된 아이тем집합 중 L_{k-1} 에 포함되지 않는 부분집합을 가지는 아이тем집합은 모두 제거한다.
 - [3.3] L_k 에 포함된 아이тем집합 중 Support 값이 c 보다 작은 아이тем집합을 모두 제거한다.
- [4] 최종 아이тем집합 리스트 $L = \bigcup_{s=1}^k L_s$ 을 결정한다.

앞의 아이тем집합 리스트 작성 알고리즘은 Support를 아이тем집합의 중요도를 평가하는 척도로 사용하고 있음을 확인할 수 있다. 또한 아이тем 구매횟수의 하한 값 c 는 시스템의 상황에 따라 외부적으로 정해지는 값이며 만약 아이тем의 구매가 굉장히 드문 경우라면 1을 사용해도 무방하다.

Step 3. 사용자 $u \in U$ 의 아이тем집합 리스트 작성

$$L_u = \{X \in L : s_{u,p} = 1, p \in X\}$$

사용자 u 의 아이тем집합 리스트 L_u 는 위에서 작성한 아이тем집합 리스트 L 에서 사용자 u 가 구매한 아이тем집합만을 선택해서 작성한 리스트이다. 만약 사용자 u 가 구매활동이 활발하여 다양한 아이тем을 구매한 경우라면 L_u 에 포함되는 아이тем집합이 적어도 하나 존재할 가능성이 매우 높다. 만약 $L_u = \emptyset$ 인 경우, 위 아이тем집합 리스트 작성 알고리즘에서 $L_k = \{X \in P : \text{supp}(X) \geq c, |X| = k\}$ 대신 사용자 u 가 구매한 리스트 $L_{u,k} = \{X \subset P : \text{supp}(X) \geq c, |X| = k, s_{u,p} = 1, p \in X\}$ 를 사용하여 새로운 리스트를 작성한다. 특별한 경우로 만약 사용자 u 가 시스템에 새로 들어온 사용자이거나 $L_{u,k}$ 를 사용하기 어려운 경우라면 $L_u = L$ 로 사용하여도 무방하다.

Step 4. L_u 에서 L 로의 연관규칙 구성

$$A = \{X \Rightarrow Y : X \in L_u, Y \in L\}$$

두 리스트 L 과 L_u 가 구성되면 아이тем집합 $X \in L_u$ 로부터 아이тем집합

$Y \in L$ 로의 연관규칙을 모두 구성할 수 있고 이를 추천 시스템에 적용할 수 있다. 예를 들어 모든 연관규칙들을 $conf(X \Rightarrow Y)$ 혹은 $lift(X \Rightarrow Y)$ 의 크기순으로 정렬하고 값이 가장 큰 아이тем집합 Y 를 사용자 u 에게 추천할 수도 있고, $conf(X \Rightarrow Y)$ 혹은 $lift(X \Rightarrow Y)$ 값의 크기의 하한을 정하여 이 값보다 큰 값을 가지는 연관규칙을 추천에 활용할 수도 있다.

구매내역 행렬 S 에서 아이тем구매횟수의 하한값 $c = 2$ 를 사용하여 아이тем p_1 을 구매한 사용자 u_1 에게 아이тем p_5 를 추천하는 과정 예시는 다음과 같다(Jannach et al., 2011).

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 & \infty \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix}, L_{u_1} = \{\{p_1\}\}$$

$$L_1 = \{\{p_1\}, \{p_3\}, \{p_4\}, \{p_5\}\}, L_2 = \{\{p_1, p_4\}, \{p_1, p_5\}\}, L_3 = \emptyset$$

$$L = L_1 \cup L_2 = \{\{p_1\}, \{p_3\}, \{p_4\}, \{p_5\}, \{p_1, p_4\}, \{p_1, p_5\}\},$$

$$conf(\{p_1\} \Rightarrow \{p_5\}) = 2/3, conf(\{p_1\} \Rightarrow Y) \leq 1/3, Y \in L$$

다. Probabilistic recommendation

Probabilistic recommendation 방법은 선호도가 주어진 아이тем집합을 활용하여 선호도가 주어지지 않은 다른 아이тем에 대한 선호도의 조건부 분포를 예측하는 방법이다. 이러한 예측 과정에서 모형을 완성하기 위한 몇 가지 조건부 확률이 필요하고 이 확률을 계산하는 과정에서 Bayes 정리와 Naive Bayes 추정법을 사용하게 된다(Breese et al., 1998; Chen and George, 1999; Miyahara and Pazzani, 2000). 먼저 Probabilistic recommendation 방법에서 가장 핵심적인 역할을 하

는 Bayes 정리는 다음과 같다.

■ R_X 가 주어진 경우의 R_p 의 조건부(conditional) 확률

$$\begin{aligned} P(R_p = x_p | R_X = x_X) \\ &= P(R_X = x_X | R_p = x_p) P(R_p = x_p) / P(R_X = x_X) \\ &= \prod_{q \in X} P(R_q = x_q | R_p = x_p) P(R_p = x_p) / P(R_X = x_X) \\ &\propto \prod_{q \in X} P(R_q = x_q | R_p = x_p) P(R_p = x_p), \end{aligned}$$

- R_p = 아이템 p 의 선호도를 나타내는 확률변수

- $R_X = (R_p)_{p \in X}$: 아이템집합 X 의 선호도를 나타내는 확률벡터

위 조건부 확률은 아이템집합 $X \subset P$ 에 대하여 x_X 의 선호도를 기록한 경우에 새로운 아이템 $p \in P$ 에 대한 조건부 확률을 어떻게 계산하는가를 제시하고 있다. 위 정의에서 첫 번째 등호는 Bayes 정리를 활용한 것이고, 두 번째 등호는 조건부 독립 조건을 추가하여 조건부 확률을 좀 더 자세하게 분해한 것이다. 또한 조건부 확률의 절대적인 값을 사용하는 경우보다 대소를 비교하여 활용하는 경우가 더 일반적이므로 고정된 확률 $P(R_X = x_X)$ 는 주어진 상수로 취급하여 제외하고 $\prod_{q \in X} P(R_q = x_q | R_p = x_p) P(R_p = x_p)$ 를 조건부 확률로 정의한다.

Step 1. 데이터의 구성

$R = (r_{u,p})_{u \in U, p \in P}$: $n \times m$ 선호도 행렬

Probabilistic recommendation 방법은 일반적으로 선호도 행렬을

대상으로 하여 진행된다. 하지만 구매내역 행렬 S 를 사용할 수 있는데 이 경우 선호도의 분포가 아니라 아이템의 구매내역에 대한 분포가 유도된다.

Step 2. 사용자 u 에 대한 아이템집합 구성

$$X_u = \{q \in P : r_{u,q} \neq \infty\}$$

위 아이템집합 X_u 는 반드시 사용자 u 의 선호도가 존재하는 아이템들만으로 구성되어야 하며, 이 정보를 활용하여 u 의 선호도가 없는 아이템 $p \notin X_u$ 에 대한 선호도를 예측하게 된다.

Step 3. Naive Bayes 추정법을 활용한 조건부 확률의 추정

$$\hat{P}(R_p = x_p) = \sum_{v \in U} I(r_{v,p} = x_p) / n, \quad p \notin X_u$$

$$\hat{P}(R_q = x_q | R_p = x_p) = \frac{\sum_{v \in U} I(r_{v,q} = x_q, r_{v,p} = x_p)}{\sum_{v \in U} I(r_{v,p} = x_p)}, \quad q \in X_u, \quad p \notin X_u$$

위 조건부 확률의 추정에서 $\hat{P}(R_p = x_p)$ 는 전체 사용자 중에서 아이템 $p \notin X_u$ 를 구매한 사용자의 비율이며, $\hat{P}(R_q = x_q | R_p = x_p)$ 는 아이템 p 를 구매한 사용자 중에서 아이템 $q \in X_u$ 를 구매한 사용자의 비율을 의미한다.

Probabilistic recommendation 방법은 선호도의 분포를 추정하는

과정이므로 앞에서 소개한 방법들에 비하여 일반적으로 계산량이 많은 편이며 데이터의 수가 충분하지 않거나 미정인 값이 너무 많이 포함된 경우 추정된 확률이 정확하지 않을 수 있다. 하지만 위에서 언급한 바와 같이 모형을 한 번 구축하면 모형의 변화가 예상되는 시점까지 계속해서 사용할 수 있으며 직관적으로 이해하기 쉽다는 장점이 있다.

선호도 행렬 R 에서 아이템집합 X 에 대한 사용자 u_1 의 선호도 벡터 x_X 를 사용하여 u_1 의 아이템 p_5 에 대한 선호도 점수가 1일 확률을 추정하는 예시는 다음과 같다(Jannach et al., 2011).

$$R = \begin{pmatrix} 1 & 3 & 3 & 2 & \infty \\ 2 & 4 & 2 & 2 & 4 \\ 1 & 3 & 3 & 5 & 1 \\ 4 & 5 & 2 & 3 & 3 \\ 1 & 1 & 5 & 2 & 1 \end{pmatrix}, X = \{p_1, p_2, p_3, p_4\}, x_X = (1, 3, 3, 2),$$

$$\hat{P}(R_{p_5} = 1) = 2/4$$

$$\hat{P}(R_{p_1} = 1 | R_{p_5} = 1) = 2/2, \hat{P}(R_{p_2} = 3 | R_{p_5} = 1) = 1/2$$

$$\hat{P}(R_{p_3} = 3 | R_{p_5} = 1) = 1/2, \hat{P}(R_{p_4} = 2 | R_{p_5} = 1) = 1/2$$

$$P(R_{p_5} = 1 | R_X = x_X) = 2/2 \times 1/2 \times 1/2 \times 1/2 \times 2/4 = 0.0625$$

4. Slope one predictors

최근 특정 아이템의 선호도를 공유하는 사용자들의 정보만을 활용하여 새로운 아이템의 선호도를 예측하는 방법들이 소개되고 있다. 이 방법들은 단순하지만 직관적인 의미를 가지며 계산이 효율적인 면에서 장점을 가진다. 특히 특정 아이템의 선호도가 다른 아이템들의 선호도에 대하여 평균적으로 얼마나 차이가 나는가를 예측하는 방법을 Slope one pre-

dictors 방법(Lemire and Maclachlan, 2005)이라고 한다. 먼저 기술에 필요한 몇 가지 용어를 다음과 같이 정의한다.

■ 아이템 $p \in P$ 와 $q \in P$ 에 대한 공통사용자 집합

$$C_{p,q} = \{v \in U : r_{v,p} \neq \infty, r_{v,q} \neq \infty\}$$

■ 아이템 $p \in P$ 의 $q \in P$ 에 대한 선호도의 편차

$$dev(p,q) = \frac{1}{|C_{p,q}|} \sum_{v \in C_{p,q}} (r_{v,p} - r_{v,q})$$

위 공통사용자 집합 $C_{p,q}$ 는 서로 다른 두 아이템 p 와 q 에 대하여 선호도가 동시에 존재하는 사용자 전체의 집합을 의미하며, 선호도의 편차 $dev(p,q)$ 는 아이템 p 에 대한 선호도가 아이템 q 에 대한 선호도에 비하여 얼마나 큰 값을 가지는지에 대하여 공통사용자 집합 $C_{p,q}$ 전체에서 평균을 계산한 것이다. Slope one predictors 방법은 선호도 행렬을 데이터로 사용하며 추천 시스템 구성의 절차는 다음과 같다.

Step 1. 사용자 $u \in U$ 의 전체 아이템집합 구성

$$X_u = \{q \in P : r_{u,q} \neq \infty\}$$

사용자 u 의 선호도 행렬을 조사하여 u 가 선호도를 작성한 기록이 존재하는 아이템들로 아이템집합 X_u 를 구성한다.

Step 2. 사용자 $u \in U$ 의 아이템 $p \in X_u^c$ 에 대한 중요 아이템집합 구성

$$X_{u,p} = \{q \in X_u : C_{p,q} \neq \emptyset\}$$

사용자 u 의 전체 아이템집합 X_u 에 속하는 모든 아이템에 대하여 새로운 아이템 $p \in X_u^c$ 와의 공통사용자 집합 $C_{p,q}$ 를 조사하고 $C_{p,q}$ 가 공집합이 아닌 아이템 q 를 모아서 사용자 u 의 아이템 p 에 대한 중요 아이템집합을 구성한다.

Step 3. 사용자 $u \in U$ 의 아이템 $p \notin X_u$ 에 대한 선호도 예측

$$pred(u,p) = \frac{\sum_{q \in X_{u,p}} |C_{p,q}| \{dev(p,q) + r_{u,q}\}}{\sum_{q \in X_{u,p}} |C_{p,q}|}$$

아이템 $p \in X_u^c$ 의 아이템 $q \in X_u$ 에 대한 선호도 편차 $dev(p,q)$ 와 선호도 $r_{u,q}$ 를 합하여 사용자 u 의 아이템 q 에 대한 선호도를 보정하고 이 보정된 선호도를 중요 아이템집합 $X_{u,p}$ 에 속하는 모든 아이템 q 에 대하여 $C_{p,q}$ 의 크기를 가중치로 하는 가중 평균을 계산하여 이 값을 u 의 p 에 대한 선호도 값으로 예측한다.

선호도 행렬 R 을 사용하여 사용자 u_1 의 아이템 p_3 에 대한 선호도를 예측하는 예시는 다음과 같다(Jannach et al., 2011).

$$R = \begin{pmatrix} 2 & 3 & \infty \\ 3 & 2 & 5 \\ 4 & \infty & 3 \end{pmatrix}, X_{u_1} = \{p_1, p_2\},$$

$$C_{p_1, p_2} = \{u_1, u_2\}, C_{p_1, p_3} = \{u_2, u_3\}, C_{p_2, p_3} = \{u_2\},$$

$$X_{u_1, p_3} = \{p_1, p_2\}$$

$$dev(p_3, p_1) = 1/2 \times ((5-3) + 3-4) = 1/2$$

$$dev(p_3, p_2) = 1/1 \times (5-2) = 3$$

$$pred(u_1, p_3) = \frac{2 \times (1/2 + 2) + 1 \times (3 + 5)}{2 + 1} = 13/3$$

제2절 Content-based recommendation

1절에서 소개한 Collaborative recommendation 방법은 아이টে에 대한 선호도 혹은 구매 내역을 가지고 추천 시스템을 구축하는 방법들이다. 이러한 방법들은 각각의 아이টে 혹은 사용자에 대한 고유의 특성을 반영하지 않기 때문에 아이টে 판매 기록이나 선호도에 대한 기록이 존재하는 대부분의 시스템에서 구성할 수 있다는 장점이 있다.

하지만 아이টে이나 사용자의 고유의 특성을 고려하여 좀 더 직접적인 방법으로 추천 시스템을 구성할 수도 있다. 예를 들어 해리 포터는 유명한 판타지 소설이고 어떤 사용자가 판타지 소설을 아주 좋아한다는 두 가지 사실을 알고 있는 경우 이 사용자에게 해리 포터를 추천하는 것이 직관적으로도 이해하기 쉽고 아이টে과 사용자의 특성을 반영한 직접적인 추천방법일 수 있다.

Content-based recommendation 방법은 이렇게 아이টে과 사용자의 고유의 특성, 예를 들면 아이টে의 카탈로그나 사용자 정보파일 등에

나타난 키워드(keyword)를 바탕으로 하여 추천 시스템을 구성하는 방법이며 시스템상에 이러한 고유의 특징이 구체적으로 기록되어 있는 경우 Collaborative recommendation 방법을 대체하여 고려할 수 있는 추천 방법이다(Charkrabarti, 2002; Adomavicius and Tuhilin, 2005; Pazzani and Billsus, 2007; Manning et al., 2008). 따라서 시스템상의 문서 혹은 문서 내에서 사용되는 키워드의 중요도나 유사도를 정의하는 기술이 Content-based recommendation 방법의 핵심을 구성한다. 이러한 특성 때문에 Content-based recommendation 방법은 Text mining 기법과 함께 소개되거나 아니면 Text mining 기법의 응용분야로 소개되기도 한다.

이 절에서는 추천 시스템 구성에 대하여 직접적으로 기술하기보다는 문서와 단어의 중요도 혹은 유사도와 같이 추천 시스템을 구축하는 과정에서 필요한 정량적인 측도를 어떻게 정의할 수 있는지에 대하여 검토하기로 한다. 기술의 편의성을 위해 다음과 같이 시스템상에 문서(document)와 단어(term) 두 가지 기준으로 분류할 수 있는 데이터가 구성되어 있다고 가정한다.

■ Content-based recommendation 방법의 데이터 index

$D = \{d_1, d_2, \dots, d_n\}$: n 개의 문서의 집합

$T = \{t_1, t_2, \dots, t_m\}$: m 개의 단어의 집합

1. TF-IDF

TF-IDF(term frequency-inverse document frequency)는 여러

개의 문서(document)가 주어졌을 때, 특정 단어(term)가 전체 문서군(set of documents)에서 얼마나 중요한 역할을 하는지에 대한 양적 수치이며(Charkrabarti, 2002; Pazzani and Billsus, 2007; Salton and Buckley, 1988; Salton et al., 1975) Text mining 방법에서 널리 사용되는 가중치 산출 방식 중의 하나이다. 따라서 전체 문서군에서 핵심이 되는 단어를 찾거나 혹은 문서 간의 유사도를 결정하는 과정에 사용된다. TF-IDF를 계산하는 절차는 다음과 같다.

Step 1. 데이터의 구성

$F = (f_{d,t})_{d \in D, t \in T} : n \times m$ 빈도(frequency) 행렬,
 - $f_{d,t}$ = 문서 $d \in D$ 에서 단어 $t \in T$ 가 사용되는 빈도

빈도행렬 F 는 빈도 $f_{d,t}$ 를 행 방향으로서는 문서가, 열 방향으로서는 단어가 위치하도록 구성한 행렬이다.

Step 2. 문서 $d \in D$ 에서 단어 $t \in T$ 의 TF(term frequency) 정의

$$TF(t, d) = \frac{1}{2} + \frac{f_{d,t}}{2 \times \max_{w \in D} f_{d,w}}$$

TF는 TF-IDF를 정의하기 위한 중간 단계에서 필요한 수치이며 단어 t 가 문서 d 에서 얼마나 자주 사용되는가를 나타낸다. TF는 다양한 방식으로 정의할 수 있지만 자주 사용되는 단어일수록 TF 값이 커지도록 정의해야 한다. 위 정의에서는 문서 d 의 길이를 고려하기 위하여 빈도 $f_{d,t}$ 를 문서에서 가장 자주 나타나는 단어의 빈도 $\max_{w \in D} f_{d,w}$ 로 나눈

비율을 최소 0.5에서 최대 1이 되도록 보정하여 정의한 것이다. 이러한 정의 이외에도 $TF(t, d) = \log(1 + f_{d,t})$ 를 사용하거나 혹은 단순히 $TF(t, d) = I(f_{d,t} \neq 0)$ 를 사용할 수도 있다.

Step 3. 문서군 D 에서 단어 $t \in T$ 의 DF(document frequency) 정의

$$DF(t, D) = \frac{1 + \sum_{d \in D} I(f_{d,t} \neq 0)}{|D|}$$

TF는 단어의 사용 빈도를 의미하므로 TF 값이 높을수록 자주 사용되는 단어, 즉 중요한 단어라고 판단할 수 있다. 하지만 특정 단어가 문서군 전체에서 지나치게 자주 사용되는 경우 이 단어는 중요한 단어가 아니라 오히려 흔한 단어일 수도 있다. DF는 이러한 상황을 고려하기 위하여 정의한 수치이며 특정 단어가 문서군 전체에서 얼마나 자주 사용되는가를 나타낸다. DF 역시 다양한 방법으로 정의할 수 있지만 DF 값이 클수록 단어의 중요도가 작아지도록 정의해야 하며 위에서 소개한 정의는 이러한 반비례 관계에서 분자 부분이 0이 되지 않도록 보정한 것이다.

Step 4. 문서군 D 에 포함된 문서 $d \in D$ 에서 단어 $t \in T$ 의 TF-IDF 정의

$$TFIDF(t, d, D) = TF(t, d) \times IDF(t, D),$$

$$IDF(t, D) = \frac{1}{DF(t, D)}$$

TF-IDF는 위에서 정의한 TF와 DF의 곱으로 정의한다. 즉 특정 문서

내에서 자주 사용되지만 전체 문서군에서는 자주 사용되지 않는 단어일 수록 TF-IDF 값이 커진다. 따라서 TF-IDF를 사용하면 전체 문서군에서 흔하게 나타나는 단어를 제거하는 효과를 얻을 수 있다.

이상의 정의로부터 계산한 TF-IDF 값을 문서가 행 방향, 단어가 열 방향이 되도록 배열한 행렬을 TF-IDF 행렬이라고 한다. 일반적으로 문서군 전체에서 사용되는 단어를 모두 고려하는 경우 TF-IDF 행렬은 크기가 매우 크고 원소의 대부분이 비어(missing) 있는 행렬이다. 따라서 중요한 단어들을 좀 더 효율적으로 구성하려면 행렬을 구성하기 전에 몇 가지 사전 작업을 진행하는 것이 좋다. 예를 들면, 특별한 의미를 가지지 않는 단어를 사전에 제거할 수도 있고(Porter, 1980) 동의어를 하나의 단어로 간주하거나 전체 단어의 수에 제한을 둘 수도 있다(Balabanovic and Shoham, 1997). 특히 특정 구문이 문서의 특징을 잘 나타낸다면 이 구문 전체를 하나의 단어로 간주할 수도 있다(Chakrabarti, 2002).

2. Probabilistic model for document classification

Content-based recommendation 방법의 핵심은 아이템의 특징과 사용자의 특징을 고려해서 일종의 맞춤형 추천을 제공하는 것이다. 따라서 사용자가 특정 문서로 정의된 아이템을 선호할 것인지 그렇지 않을 것인지를 결정하는 것이 추천과정에서 매우 중요한 역할을 하게 된다. Probabilistic model 방법은 이러한 분류의 문제를 해결하기 위한 확률 모형을 결정하는 방법이다. 따라서 Probabilistic model 방법을 적용하기 위해서는 단어와 문서뿐만 아니라 사용자들의 문서에 대한 선호도 혹은 중요도가 사전에 조사되어야 한다.

가. Boolean feature vector classification

Boolean feature vector classification 방법은 각각의 문서에서 각각의 단어가 사용되는 패턴을 사용하여 문서가 중요한 문서일 확률을 추정하는 방법이다(Pazzani and Billsus, 1997; Friedman, 1997; McCallum and Nigam, 1998). 특정 문서에 특정 단어가 사용되면 1 그렇지 않으면 0으로 작성된 이진 데이터 행렬(Boolean design matrix)과, 각 문서가 중요하면 1 그렇지 않으면 0으로 작성된 이진 중요도 벡터(Boolean importance vector)를 데이터로 사용하며 이 데이터로부터 아래의 Bayes 정리를 적용하여 중요도에 대한 조건부 확률을 추정하는 방법이다.

■ R_T 가 주어진 경우의 S_d 의 조건부(conditional) 확률

$$\begin{aligned} P(S_d = s_d | R_T = x_T) \\ &= P(R_T = x_T | S_d = s_d) P(S_d = s_d) / P(R_T = x_T) \\ &= \prod_{t \in T} P(R_t = x_t | S_d = s_d) P(S_d = s_d) / P(R_T = x_T) \\ &\propto \prod_{t \in T} P(R_t = x_t | S_d = s_d) P(S_d = s_d), \end{aligned}$$

- S_d = 문서 $d \in D$ 의 중요도를 나타내는 이진 확률변수
- R_t = 단어 $t \in T$ 의 사용 여부를 나타내는 이진 확률변수
- $R_T = (R_t)_{t \in T}$: 전체 단어의 사용 여부를 나타내는 이진 확률벡터

위 조건부 확률은 전체 단어가 문서에서 사용되고 있는지를 이진 벡터 x_T 로 기록한 경우에 문서의 중요도에 대한 조건부 확률을 어떻게 계산하는가를 보여 주고 있다. 위 정의에서 첫 번째 등호는 Bayes 정리를 활용

한 것이고, 두 번째 등호는 조건부 독립 조건을 추가하여 조건부 확률을 좀 더 자세하게 분해한 것이다. 또한 조건부 확률의 절대적인 값을 사용하는 경우보다 대소를 비교하여 활용하는 경우가 더 일반적이므로 고정된 확률 $P(R_T = x_T)$ 는 주어진 상수로 취급하여 제외한다.

Step 1. 데이터의 구성

$B = (b_{t,d})_{t \in T, d \in D}$: $n \times m$ 이진(Boolean) 빈도(frequency) 행렬,

- $b_{d,t}$ = 문서 $d \in D$ 에서 단어 $t \in T$ 가 사용되면 1, 아니면 0임

$L = (\ell_d)_{d \in D}$: 길이가 n 인 중요도 벡터,

- ℓ_d = 문서 $d \in D$ 가 중요하면 1, 아니면 0임

Step 2. Naive Bayes 추정법을 활용한 조건부 확률의 추정

$$\hat{P}(S_d = s_d) = \frac{1}{n} \sum_{d \in D} I(\ell_d = s_d)$$

$$\hat{P}(R_t = x_t | S_d = s_d) = \frac{\sum_{d \in D} I(b_{d,t} = x_t, \ell_d = s_d)}{\sum_{d \in D} I(\ell_d = s_d)}$$

위 두 조건부 확률의 추정량을 Bayes 정리에 적용하면 문서 d 의 중요도를 나타내는 사후확률 $\hat{P}(S_d = 1 | R_T = x_T)$ 을 계산할 수 있고 이 값이 0.5보다 크면 문서 d 를 중요한 문서로 분류한다. 참고로 전체 문서군에서 중요한 문서의 개수가 매우 적은 경우라면 0.5보다 작은 적당한 값 $c < 0.5$ 를 사전에 정의하여 이 값을 분류의 기준으로 정하는 것이 좀 더

합리적이다.(박창이 등, 2015)

다음은 이진 빈도행렬 B 와 이진 중요도 벡터 L 을 사용하여 단어 사용 여부가 x_T 로 기록된 문서 d 의 중요도에 대한 확률을 추정하는 예제이다 (Jannach et al., 2011).

$$B = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, L = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, x_T = (1, 1, 0, 0),$$

$$\hat{P}(S_d = 1) = 3/5$$

$$\hat{P}(R_{t_1} = 1 | S_d = 1) = 3/3, \hat{P}(R_{t_2} = 1 | S_d = 1) = 2/3$$

$$\hat{P}(R_{t_3} = 0 | S_d = 1) = 1/3, \hat{P}(R_{t_4} = 0 | S_d = 1) = 2/3$$

$$\hat{P}(S_d = 1 | R_T = x_T) \propto 3/3 \times 2/3 \times 1/3 \times 2/3 \times 3/5 = 0.088$$

나. Term counts classification

Boolean feature vector classification 방법은 각각의 단어가 각각의 문서에 사용되었는지 그렇지 않은지만 조사하면 되기 때문에 데이터의 구성이 매우 쉽다는 장점을 가진다. 하지만 오히려 단어의 중요도나 빈도가 전혀 반영되지 않는다는 단점으로 이해할 수도 있다. 이러한 문제를 해결하기 위해서는 위에서 단어의 중요도를 사전에 정의하고 이러한 중요도를 확률추정 과정에서 가중치로 사용하는 등의 보정 작업을 거쳐야 한다. Term counts classification 방법은 이러한 단점을 보완하는 단순한 방법의 하나로 각 단어의 중요도를 그 단어의 빈도로 사용하는 방법이다(Manning et al., 2008).

Step 1. 데이터의 구성

$F = (f_{d,t})_{d \in D, t \in T}$: $n \times m$ 빈도(frequency) 행렬

- $f_{d,t}$ = 문서 $d \in D$ 에서 단어 $t \in T$ 가 사용되는 빈도

$L = (\ell_d)_{d \in D}$: 길이가 n 인 중요도 벡터

- ℓ_d = 문서 $d \in D$ 가 중요하면 1, 아니면 0임

위 빈도행렬 F 는 TF-IDF를 정의하는 과정에서 정의한 빈도행렬과 동일하며, 중요도 벡터 L 은 Boolean feature vector classification 방법에서 정의한 벡터이다. 아래 Bayes 정리는 특정 단어의 빈도가 주어진 경우에 문서의 중요도에 대한 조건부 확률을 계산하는 방법을 보여 준다.

■ R_T 가 주어진 경우의 S_d 의 조건부(conditional) 확률

$$\begin{aligned} P(S_d = s_d | R_T = x_T) \\ &= P(R_T = x_T | S_d = s_d) P(S_d = s_d) / P(R_T = x_T) \\ &= \prod_{t \in T} P(R_t = x_t | S_d = s_d) P(S_d = s_d) / P(R_T = x_T) \\ &\propto \prod_{t \in T} P(R_t = x_t | S_d = s_d) P(S_d = s_d), \end{aligned}$$

- S_d = 문서 $d \in D$ 의 중요도를 나타내는 이진 확률변수

- R_t = 단어 $t \in T$ 의 빈도를 나타내는 확률변수

- $R_T = (R_t)_{t \in T}$: 전체 단어의 빈도를 나타내는 확률벡터

Step 2. Term counts를 고려한 Naive Bayes 추정법

$$\hat{P}(S_d = 1) = \frac{1}{n} \sum_{d \in D} I(\ell_d = 1)$$

$$\hat{P}(R_t = x_t | S_d = 1) = \left\{ \frac{1 + \sum_{d \in D} f_{d,t}}{|T| + \sum_{d \in D} \sum_{s \in T} f_{d,s} I(\ell_s = 1)} \right\}^{x_t}$$

위 확률의 추정과정에서 $\hat{P}(S_d = 1)$ 은 전체 문서군에서 중요한 문서들의 비율을 의미하며 $\hat{P}(R_t = x_t | S_d = 1)$ 는 중요한 문서에서 단어 t 가 x_t 번 사용될 확률을 의미한다. $\hat{P}(R_t = x_t | S_d = 1)$ 는 전체 문서군에서 단어 t 가 사용된 빈도 $\sum_{d \in D} f_{d,t}$ 와 중요한 문서에서 사용된 모든 단어의 개수 $\sum_{d \in D} \sum_{s \in T} f_{d,s} I(\ell_s = 1)$ 의 비율에 대한 x_t 제곱으로 정의되며 $|T|$ 는 추정확률이 전체 단어의 종류에 대한 정보를 반영하도록 보정하는 항이다.

빈도 행렬 F 와 중요도 벡터 L 을 사용하여 단어의 사용 여부가 x_T 로 기록된 문서 d 의 중요도에 대한 조건부 확률은 다음과 같이 계산될 수 있다(Jannach et al., 2011).

$$F = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \quad L = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad x_T = (3, 0, 0, 0, 1, 1),$$

$$|T| = 6, \quad \hat{P}(S_d = 1) = 3/4$$

$$\hat{P}(R_{t_1} = 3 | S_d = 1) = \{(1+5)/(6+8)\}^3$$

$$\hat{P}(R_{t_5} = 1 | S_d = 1) = \{(1+0)/(6+8)\}^1$$

$$\hat{P}(R_{t_5} = 1 | S_d = 1) = \{(1 + 0) / (6 + 8)\}^1$$

$$P(S_d = 1 | R_T = x_T) \propto 3/4 \times (3/7)^3 \times 1/14 \times 1/14 = 0.0003$$

다. Other classification methods

위 Probabilistic classification 방법에서 사용된 데이터 구조는 Machine learning 방법에서 사용되는 분류방법에 바로 적용할 수 있다. 예를 들면, B 와 F 를 데이터행렬(design matrix) L 을 분류변수 벡터(target vector)로 간주하면 바로 Logistic regression model 방법(Nelder and Wedderburn, 1972)을 적용할 수 있다. 물론 TF-IDF에서 제공하는 가중치를 활용하면 Weighted logistic regression model 방법을 적용할 수도 있다. 이 밖에도 Support vector machine(Joachims, 1998), Decision tree(Quinlan, 1993; Kim et al., 2001), Neural network(Bishop, 1995) 방법 등 다양한 방법을 적용할 수 있다.

라. Feature selection

앞에서 잠깐 언급하였듯이 행렬 B 와 F 는 크기가 매우 크고 원소의 대부분이 비어(missing) 있는 행렬이다. TF-IDF 단계에서 이러한 문제를 해결하기 위한 몇 가지 직관적인 방법을 소개했지만 중요도 벡터 L 이 있는 경우에는 좀 더 이론적인 방법을 적용할 수 있다. 이러한 방법 중 가장 대표적인 방법이 χ^2 -통계량을 사용하는 방법이다.

■ 단어 $t \in T$ 의 χ^2 -통계량

$$\chi^2(t) = \frac{(A_t + B_t + C_t + D_t)(A_t D_t - B_t C_t)^2}{(A_t + B_t)(A_t + C_t)(B_t + D_t)(C_t + D_t)}$$

$$- A_t = \sum_{d \in D} I(b_{d,t} = 1, \ell_d = 1), B_t = \sum_{d \in D} I(b_{d,t} = 0, \ell_d = 1),$$

$$- C_t = \sum_{d \in D} I(b_{d,t} = 1, \ell_d = 0), D_t = \sum_{d \in D} I(b_{d,t} = 0, \ell_d = 0).$$

위 χ^2 -통계량은 단어 t 의 사용 여부와 문서의 중요도에 대한 독립성 검정에 사용되는 통계량이다. 따라서 단어 t 의 χ^2 -통계량이 작을수록 t 가 문서의 중요도에 미치는 영향이 크다고 할 수 있다. 이러한 성질을 바탕으로 하여 모든 단어에 대하여 χ^2 -통계량을 계산한 후에 작은 순으로 나열하여 상위 몇 개의 단어를 사용하면 전체 데이터를 사용하지 않아도 되는 장점이 있다.

제3절 소결

이 장에서는 추천 시스템(Recommendation system)의 방법론과 알고리즘 및 최신 동향을 살펴보았다. 추천 시스템의 이론적 검토 및 최신 동향 파악은 보건복지데이터포털에 적용할 수 있는 시스템의 수준을 정하고, 장기적으로 어떠한 정보를 수집해야 하는지 예상하여 고도화할 수 있는 확장 가능한 프레임워크를 준비하는 데 필요한 부분이다. 추천 시스템은 네이버(주)나 Amazon과 같은 기업들의 핵심기술로, 다양한 방법론과 모형을 적합하여 이용자들에게 다양한 서비스를 제공해 주는 매우

중요한 역할을 하고 있다. 2015년 말, 네이버(주)에서는 선물 추천 의도를 가진 검색어를 입력 할 경우 성별, 연령별로 구매할 만한 상품 정보를 자동으로 추천해 주는 ‘선물추천 검색’을 도입하였다. 이는 이용자의 클릭, 구매 내력 등의 정보를 이용하여 User's preference analysis, Rich content service 등의 기술로 검색 의도에 따라 구매할 만한 상품들을 자동으로 추출하여 추천하는 시스템이다. 공공기관에서도 추천 시스템과 관련하여 시도는 하고 있으나, 아직은 도입단계여서 한정된 자료로 서비스를 제공하고 있는 실정이다. 보건복지데이터포털에서도 올해 처음으로 보건복지데이터포털의 이용자들에게 맞춤형 서비스를 제공하고자, 이용자가 원하는 자료를 찾았던 유저들의 패턴을 분석하여 비슷한 자료를 추천해 주는 시스템을 도입하였다. 위의 이론적 검토를 바탕으로 하여, 보건복지데이터포털에서의 적용 가능성 및 고도화 방향을 살펴보면 다음과 같다. 현재는 보건복지데이터포털 이용자의 로그정보를 활용하여 선호도를 예측하여 추천해 주는 collaborative recommendation system의 기초적인 추천기능이다. Association rule mining 방법은 보건복지데이터포털의 Data bank에서 데이터를 구매한 사용자들의 구매 패턴 간의 연관규칙을 추출하여 추천에 도입할 수 있다. 보건복지데이터포털에서 제공하는 데이터의 수가 많아질수록 아이템(데이터)집합리스트 작성 알고리즘은 중요해진다. 보건복지데이터포털에 새로 들어온 사용자인 경우, $L_u = L$ 로 정의하여 적용한다. Term counts classification 방법은 Content-based recommendation으로 단어의 중요도나 빈도를 반영하여 사전에 정의하고, 이러한 중요도를 확률추정 과정에 가중치로 사용할 수 있다. 이는 연구DB에서 연구보고서를 추천할 때, 연구보고서의 내용을 단어 빈도와 중요성으로 측정하여 고려할 수 있는 방법이다. 그리고 이용자가 보건복지데이터포털의 웹서비스를 처음 이용할 경우, 관심분야

를 설정하여 이를 이용자 PC 쿠키로 저장하게 하여 관심사를 반영한 추천도 생각해 볼 수 있다. 또한, 현재 주제별 데이터 페이지에서 하는 추천은 임의적인 구분을 해 놓은 주제와 관련된 추천이지만, 이를 확장하여 이용자별로 실제 이용하는 페이지들에서 도출되는 주제와 관련된 추천기능도 설정할 수 있을 것이다. 추천 시스템은 일회적으로 구축하고 끝나는 것이 아니라, 단계적으로 확장하여 이용자들에게 향상된 시스템으로 지속적인 서비스를 제공해야 할 것이다. 앞서 언급한 보건복지데이터포털의 적용 가능성과 고도화 방향은 내년의 추천 시스템 고도화에 고려하여 적용할 예정이다.

제 5 장

보건복지데이터포털 고도화 방향: 마이크로데이터 연계

제1절 데이터 연계방안에 대한 고찰

제2절 데이터 연계 평가 및 쟁점

5

보건복지데이터포털 고도화 << 방향: 마이크로데이터 연계

제1절 데이터 연계방안에 대한 고찰

과학적 의사결정을 위한 전통적인 데이터 수집 방법으로 사용되던 실험과 조사자료는 빅데이터 시대의 도래로 다소 그 관심이 줄어들고 있으나 인과관계(causal relationship) 분석을 위한 실험 설계, 그리고 유한 모집단을 구성하는 개체들(elements)의 실태 내지 인식을 나타내는 모수를 통계적으로 타당한 절차에 의하여 추정하기 위한 조사(survey)는 여전히 매우 중요한 데이터수집 도구이다. 국내에서 정부 정책에 필요한 기초 자료로 사용되는 정부 승인 통계는 2016년 5월 기준으로 총 955종이며 이 중 조사통계가 418종으로 약 44%를 차지하고 있어 여전히 조사자료는 국가통계 작성에서도 중요한 역할을 차지하고 있다.

그러나 개인정보보호에 대한 인식 변화, 1인 가구의 증가, Information and Communications Technologies(ICT) 기술의 발달 등은 조사환경의 악화를 가져왔고 이는 조사비용의 증가 및 조사 데이터의 신뢰도 하락의 큰 요인이 되었다. 이는 빅데이터에 대한 관심과 더불어 데이터 수집 비용과 생산되는 통계의 신뢰도 측면에서 효율적인 통계 생산 방안에 대한 수요를 증가시켰고 이와 관련하여 여러 가지 방안들이 최근에 모색되고 있다.

특별한 목적을 가지고 과학적인 설계 및 과정을 통해 수집되는 실험 데이터 혹은 조사 데이터와는 달리 최근 민간분야에서는 무선전화 기록, 위치 정보 혹은 신용카드 기록 등을 활용한 다양한 빅데이터 분석 방안들

이 제시되고 있다. 대표적인 빅데이터 분석 방안으로는 머신러닝(machine learning), 딥러닝(deep learning), 인공 신경망 분석(artificial neural network), 데이터 마이닝(data mining) 등이 있다.

민간분야의 데이터 수집 및 이의 분석 목적과는 다르게 공공분야에서는 분석자의 측면보다는 오히려 데이터 공급자로서의 역할이 우선 고려되어야 한다. 즉 데이터 공급자로서 사용자들이 신뢰할 수 있는 데이터를 수집하고 이를 제공하는 것이 공공분야에서 데이터 수집의 중요한 목적이 된다. 이러한 측면에서 분석 방법에 초점을 맞춘 민간분야의 빅데이터 연구와는 달리 공공분야의 빅데이터 개념은 최근 관심을 받기 시작한 데이터 증대(또는 데이터 증가, data augmentation)에 가깝다고 볼 수 있다.

데이터 증대란 기존 데이터들을 적절한 방법을 통해 보완하거나 결합하여 더 다양한 형태로 분석할 수 있도록 데이터의 규모를 키우는 것을 의미한다. 대표적인 통계적 데이터 증대 방법으로는 무응답 대체(imputation), 데이터 연계(data linkage), 데이터 매칭(data matching)이 있다. 결측값을 적절한 수치로 채우는 무응답 대체는 일종의 데이터 보완으로 이해할 수 있으며, 데이터 연계나 데이터 매칭은 여러 데이터의 결합으로 이해할 수 있다. 데이터 연계와 데이터 매칭은 실제 혼동되어서 사용되는 경우가 많은데 본 절에서는 데이터 매칭을 데이터 연계의 한 방안으로 간주한다.

대표적인 국내의 데이터 연계 사례는 5년마다 수행되던 기존의 센서스 방식 대신 정부기관에서 보유하고 있는 행정데이터를 이용하여 매해 공식적인 인구, 가구 및 주택에 대한 모집단 통계를 생성하기 위하여 실시한 2015년 등록 센서스이다. 등록 센서스를 위해서 13개 행정기관으로부터 수집된 24종의 데이터가 연계되어 활용되었다. 실제 24종의 데이터

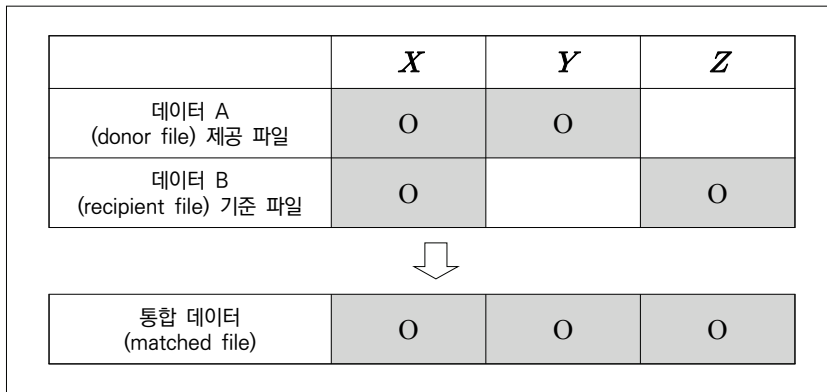
의 연계 및 사용을 위해서 데이터들의 표준화 작업과 에디팅(editing)을 포함한 많은 사전 작업 절차가 수행되었다.

본 절에서는 공공 데이터를 활용하기 위한 방안으로 데이터 증대 방안에 대한 내용들을 살펴본다. 실제 보고를 통해 수집된 공공데이터는 그 성격에 따라 주민등록번호 혹은 사업체 등록번호와 같은 고유 인식코드가 있는 경우가 있으며 이를 활용한 데이터 연계가 이루어질 수 있을 것이다. 그러나 대부분의 조사 데이터의 경우에는 서로 다른 조사를 위한 표본이 중복되거나 일치하지 않기 때문에 설령 각 개체의 고유 인식번호가 존재하는 경우에도 직접적인 데이터 연계는 가능하지 않다. 이러한 경우에 고려되는 방안이 통계적 매칭이며 이에 대한 세부적인 내용들도 살펴보고자 한다.

데이터 연계의 문제를 단순하게 살펴보기 위하여 두 데이터를 연계하는 경우를 고려한다. 오미에 외(2014)에 따르면, [그림 5-1]과 같이 서로 다른 두 개의 데이터가 존재하며 파일 A는 변수 X, Y로, 파일 B는 변수 Y, Z로 구성되어 있는 상황에서 파일 B에 파일 A를 연계한다고 가정하면, 파일 B는 기준이 되는 기존파일(host file) 또는 수용파일(recipient file)이라고 하며 본 연구에서는 ‘기준 파일’로 정의한다. 파일 A는 통합 연계과정에서 추가적인 정보를 얻기 위해 사용될 파일로 공여파일(donor file)이라고 하며, 여기서는 ‘제공 파일’이라고 정의한다. 파일 A와 파일 B에서 모두 관찰되는 변수 X는 공통변수(common variable)이며, 파일 A에서만 관찰되는 변수 Y와 파일 B에서만 관찰되는 변수 Z를 고유변수(unique variable)라고 한다. 일반적으로 데이터 매칭은 공통 변수를 이용하여 파일 B에 있는 변수 Z를 파일 A에 추가하는 것으로, 이렇게 생성된 파일을 결합파일(matched file)이라고 하며 여기서는 연계된 결과를 의미하는 통합파일이라고 정의한다. 흔히 [그림 5-1]과 같이

두 데이터를 연계하는 방안을 마이크로데이터 연계 혹은 마이크로데이터 매칭으로 명명한다. 대응되는 용어인 매크로 매칭은 두 데이터를 생산하는 확률 시스템 혹은 확률분포를 가정하고 이를 정의하기 위해 필요한 모수 추정을 위해 두 데이터를 사용하는 방안을 의미하며, 본 장에서는 기본적으로 마이크로데이터 연계에 대한 내용을 다룬다.

[그림 5-1] 자료연계의 기본구조



주: 이영섭 등. (2009). 통계조사자료와 행정자료 간의 통계적 매칭기법에 관한 연구, 통계 연구, 제 14권 제1호, pp.82-98 참고, 오미애 등. (2014). 보건복지통계정보 생산 및 활용 촉진을 위한 마이크로데이터 통합 연계 방안. 한국보건사회연구원 재구성.

마이크로데이터 연계에 대한 기존 연구는 크게 연계 혹은 매칭 알고리즘에 대한 연구와 연계 및 통합 기법에 대한 연구로 구분할 수 있다. 대표적인 알고리즘 연구로는 단계적 매칭 알고리즘(Van Pelt 2001), k-최근접 이웃 매칭 알고리즘(Van der Putten et al., 2002), 회귀분석 매칭 알고리즘(Ingram et al., 2000), 회귀분석과 k-최근접 이웃 방법의 결합 매칭 알고리즘(정성석 등, 2004), 랜덤 핫덱 방법 등이 있다. 매칭 기법에 대한 주요한 연구 결과와 매칭 기법별 근사성의 측정 방법과 매칭 과정 등 자료 연계를 위한 통계적 매칭 알고리즘 연구로는 통계개발원의 연구보고

서(2013)가, 국내의 연계 및 통합 기법에 대한 방법론 및 사례연구로는 정성석 등(2004), 고은애(2004), 안일호 (2003), 이병희(2008)가 있다.

1. 데이터 연계를 위한 이론적 가정

앞서 살펴본 바와 같이 마이크로데이터 연계는 서로 다른 자료들을 하나의 마이크로 자료로 결합하는 방안을 의미한다. 이는 두 데이터에 공통적으로 주어진 변수 벡터 X 가 포함한 정보에 따라 각 개체의 고유 인식코드를 이용하는 정확 연계(exact linkage)와, X 에 근거하여 서로 가장 유사한 개체를 연계하는 통계적 매칭(statistical matching)으로 구분할 수 있다(Moriarity, 2009).

정확연계를 위해서는 두 데이터에 모두 포함된 개체의 비율이 일정 수준 이상이어야 하며, 이들을 식별할 수 있는 고유코드가 존재해야 한다. 또한 두 데이터에 공통적으로 관측된 X 변수들에 대한 표준화 작업도 이루어져야 한다. 만약 두 데이터에서 동일한 개체로부터 측정된 변수의 값이 일치하지 않는 경우에는 두 데이터의 생성과정에서 발생한 측정오차(measurement error)에 대한 분석이 필요하고 이에 근거하여 연계된 자료의 X 값을 조정해야 한다.

마이크로데이터 연계의 목적은 증대된 자료를 이용하여 더 다양한 모집단을 분석하는 것에 있다. [그림 5-1]을 살펴보면 자료 A를 이용해서는 (X, Y) 의 모집단 분포에 대한 추론이 가능하며, 자료 B를 이용해서는 (X, Z) 의 모집단 분포에 대한 추론이 가능하나 데이터 A 혹은 B만을 이용하여 (X, Y, Z) 혹은 (Y, Z) 의 분포에 대한 추론은 가능하지 않다. 이러한 기존 마이크로데이터의 분석적 측면에서의 한계를 극복하기 위하여 데이터 증대의 한 방안으로 데이터 연계가 고려될 수 있다. 두 마이크로

데이터를 이용한 정확 연계를 통해 생성된 자료는 그 특성상 연계 후 (X, Y, Z) 의 값이 동일한 개체로부터 측정되기 때문에 (X, Y, Z) 혹은 (Y, Z) 의 모집단 분포에 대한 직접적인 추론이 가능하다.

그러나 두 마이크로데이터에 있는 유사한 개체를 연계하는 통계적 매칭을 위해서는 정확 연계 방안에서와는 달리 추가적인 가정들이 매칭을 위해서뿐만 아니라 연계된 데이터의 분석을 위해서도 필요하다. 다만, 이러한 가정들은 데이터 연계를 위해서는 필수적인 요소이지만 분석의 측면에서는 분석 가능한 방안을 제한하는 요소일 수 있다.

데이터 연계를 위해 요구되는 첫 번째 가정은 연계하고자 하는 마이크로데이터들은 동일한 모집단을 대표하는 표본들로부터 생성되어야 한다는 것으로, 연계가 이루어진 자료를 이용하여 모집단 특성에 대한 추정량의 일치성을 유지하기 위하여 필요한 가정이다. 필요한 경우에는 연계 전에 각 데이터를 적절하게 조정함으로써 동일 모집단에 대한 가정을 만족시킬 수도 있다.

두 번째 가정은 데이터 연계에서 가장 중요한 가정으로 공통 변수 X 가 주어졌을 때, 두 조사의 고유변수인 Y 와 Z 사이에 다음과 같은 조건부 독립관계가 성립되어야 한다는 것이다. 이를 조건부 독립성 가정(CIA: conditionally independent assumption)이라고 하며 다음과 같이 표현할 수 있다.

$$P(Y, Z | X) = P(Y | X) \times P(Z | X) \quad (1)$$

유한모집단이나 이산 확률분포(discrete probability distribution) 분석을 하는 경우에도 동일하게 다음의 논의가 적용되거나 편의를 위해 연속형 변수의 무한모집단 분포의 추론을 가정한다. 조건부 독립성의 가정

하에서 (X, Y, Z) 의 결합 확률밀도함수, $f(x, y, z)$ 는 다음과 같이 표현할 수 있다.

$$f(x, y, z) = f_{Y|X}(y|x)f_{Z|X}(z|x)f_X(x) \quad (2)$$

여기서 $f_{Y|X}(y|x)$ 와 $f_{Z|X}(z|x)$ 는 각각 $X=x$ 로 주어진 경우의 Y 와 Z 의 조건부 확률밀도함수를 나타내며 $f_X(x)$ 는 X 의 주변 확률밀도함수를 나타낸다. 식 (2)는 조건부 독립성 가정을 통해 $X=x$ 로 주어진 경우에 Y 와 Z 의 조건부 결합 확률함수를 아래와 같이 표현하게 된다.

$$f(x, y|z) = f_{Y|X}(y|x)f_{Z|X}(z|x)$$

즉, 특별한 가정 없이는 분석이 가능하지 않았던 (X, Y, Z) 혹은 (Y, Z) 의 모집단 결합 분포에 대한 추론이 조건부 독립성 가정을 통해서 가능하게 된다. 예를 들어, (X, Y, Z) 혹은 (Y, Z) 의 결합분포는 식 (2)를 통해 각각 다음의 식을 통해 추정할 수 있음을 알 수 있다.

$$\hat{f}(x, y, z) = \hat{f}_{A, Y|X}(y|x)\hat{f}_{B, Z|X}(z|x)\hat{f}_{AB, X}(x) \quad (3)$$

$$\hat{f}(y, z) = \int_x \hat{f}(x, y, z)dx \quad (4)$$

위 식에서 $\hat{f}(\cdot)$ 은 추정된 확률밀도함수로 첨자 A 와 B 그리고 AB 는 각각 데이터 A 와 B 그리고 AB 를 이용해서 함수 f 가 추정되었음을 나타낸다.

유한모집단의 추론이 주 관심이 되는 전통적인 조사자료 통계학 분야에서는 (X, Y, Z) 혹은 (Y, Z) 의 모집단 분포 혹은 변수들 간의 연관관계를 파악하기 위하여 연계된 자료를 직접 이용하는 방안도 고려할 수 있으며, 실제 조건부 독립의 가정이 성립하는 경우에는 식 (3)과 (4)를 통한 추론의 결과와 연계된 자료를 이용하여 산출되는 추론의 결과는 일치하게 된다. 실제 R을 포함한 많은 통계 패키지에서는 조건부 독립성 가정 하에서 통계적 매칭을 수행할 수 있는 알고리즘을 제공하고 있다.

2. 데이터 연계의 종류

각 개체의 고유 식별 변수를 이용한 정확 연계를 제외하고 조건부 독립의 가정하에서 수행할 수 있는 데이터 연계의 종류는 데이터 연계의 목적 및 연계 방법에 따라 분류할 수 있다. 우선 데이터 연계의 목적에 따라, 앞에서 언급한 결합 확률밀도함수와 같이 모집단의 분포를 정의하기 위한 모수를 추정하는 것이 목적인 매크로 연계와, 데이터 사용자에게 서로 다른 조사로부터 수집된 정보를 결합하여 보다 다양한 통계 분석이 가능하도록 결합 데이터를 제공하는 것을 목적으로 하는 마이크로데이터 연계로 구분할 수 있다.

또한 모수 추정과 자료 결합을 위한 통계 모형의 가정 여부에 따라 모수적 방법과 비모수적 방법으로 구분할 수 있다. 두 방법 모두 매크로 및 마이크로데이터 연계를 위해 사용할 수 있으며 마이크로데이터 연계를 위해서는 두 방안의 혼합방식도 고려할 수 있다.

여기에서는 마이크로데이터 연계를 중심으로 하여 연계 가능한 방안들을 살펴본다. 데이터 연계를 통한 결합 마이크로데이터 생성은 무응답 대체 과정과 상당히 유사하다. 즉 각 자료에 존재하는 항목 무응답을 적

절한 값으로 대체하는 과정으로 마이크로데이터 연계과정을 이해할 수 있다.

논의의 전개를 위해 두 파일 A 와 B 를 각각 제공(donor) 파일과 기준(recipient) 파일로 정의하자. 이 경우 데이터 연계는 B 에 존재하지 않는 Y 의 값을 파일 A 로부터 적절한 기법을 통해 추출하여 채우는 대체과정으로 이해할 수 있다. 가장 보편적인 방안은 통계적으로 타당한 기준에 따라 기준 파일과 제공 파일에 모두 존재하는 X 를 기준으로 하여 기준 파일에 있는 각 개체와 가장 유사한 개체를 제공 파일에서 선택하고 선택된 제공의 Y 의 값 혹은 이에 랜덤 오차를 추가한 값으로 기준 파일의 Y 값을 대체하는 방법이 사용된다.

이 과정에서 제공 파일에 존재하는 개체의 중복 사용 가능 여부에 따라 마이크로데이터 매칭 방안은 비제한적 방법(unconstrained matching)과 제한적 방법(constrained matching)으로 구분된다. 비제한적 연계 방법은 두 자료를 연계할 때 제공 자료를 복원 추출로 연계하는 방법으로, 동일한 개체가 중복으로 자료 연계를 위해 사용될 수 있는 방법이다. 반면에 제한적 매칭의 경우에는 제공 파일의 각 개체는 연계를 위해 한 번씩만 사용될 수 있다. 이를 위해 일반적으로 제공 파일이 기준 파일보다 규모가 큰 것으로 가정한다(D'Orazio et al., 2006).

제한적 및 비제한적 연계가 모두 적용될 수 있는 여러 연계 방안들을 모수적 방안, 비모수적 방안, 혼합적 방안으로 구분하고 각 방안별로 대표적인 연계 방안을 살펴보면 다음과 같다.

가. 모수적 마이크로데이터 연계 방안

모수적 방법을 이용한 마이크로데이터 연계 방안은 관측되지 않은 결

측 자료를 특정한 통계 모형의 가정하에서 예측한 값으로 대체하는 방안으로 흔히 예측적 접근(predictive approach)이라고도 한다. 모수적 방법을 이용한 마이크로데이터 연계를 위해서는 조건부 평균 매칭(conditional mean matching)과 조건부 예측 분포에 근거한 랜덤 추출(draws based on a conditional predictive distribution)의 두 가지 방법이 널리 사용된다.

1) 조건부 평균 매칭

모수적 방법을 이용한 마이크로데이터 연계의 핵심은 개별적인 결측 자료를 관측된 자료들로부터 계산된 결측 변수의 기댓값으로 대체하는 것이다. 즉, 이 방법은 파일 A와 B에서 미관측된 결측 자료를 다음의 조건을 만족하는 조건부 기댓값의 추정치로 대체하는 것을 의미한다.

데이터 A에서 결측된 Z 값의 대체: $\tilde{z}_a = E(Z|X=x_a)$

데이터 B에서 결측된 Y 값의 대체: $\tilde{y}_b = E(Y|X=x_b)$

일반적으로 조건부 기댓값의 선형근사추정치는 X 를 설명변수로 사용하는 회귀모형을 통해 구할 수 있다. 즉, 각 데이터를 이용하여 Y 와 Z 를 종속변수로 하는 회귀모형을 적합한 후, 이를 이용하여 각 자료에 발생한 결측값을 대체하게 된다. 특별히 연속형의 관측 자료들이 정규분포를 따를 때는 회귀식을 통해 계산되는 예측치와 조건부 기댓값은 일치하게 된다.

조건부 평균 매칭 방법은 결측 자료를 각 변수의 기댓값으로 대체하기 때문에 Y 와 Z 의 결측값에 이의 조건부 기댓값이 반복적으로 대체되

고, 이로 인하여 일반적인 평균 대체법이 갖는 분포 왜곡의 위험성을 내포한다.

이러한 조건부 평균 매칭 방안의 단점을 보완하기 위하여 추정된 회귀식과 더불어 잔차의 분산을 추정하고 이를 이용하여 백색잡음(white noise)을 생성하고 이를 예측값에 추가하는 방안이 고려된다. 예를 들어, 데이터 A에서 결측된 Z 값의 대체를 위하여서 $\tilde{z}_a = E(Z|X=x_a)$ 대신 아래의 값을 고려하게 된다.

$$\tilde{z}_a + \hat{e}, \quad \hat{e} \sim N(0, \hat{\sigma}^2)$$

2) 조건부 예측 분포에 근거한 연계

언급한 바와 같이 조건부 평균 매칭 방법은 연계를 위해 사용되는 X 의 값이 동일한 개체에 대해서는 동일한 값이 대체되어 연계된 자료에 기반을 둔 Y 와 Z 의 주변분포 및 결합분포는 왜곡될 수 있는 위험이 존재한다.

위에서 소개한 잔차를 생성하여 이러한 단점을 극복하는 방안과 더불어서 직접적으로 예측을 위한 모수 모형을 예측하고, 예측된 모형으로부터 결측을 채우기 위한 값을 무작위로 추출하는 방안을 고려할 수 있다. 즉, 조건부 독립 가정하에서 데이터 A의 결측 자료 Z 는 $f_{Z|X}(z|x_a; \hat{\theta}_{z|x})$ 분포로부터 무작위로 추출하고, 조사 B의 결측 자료 Y 는 $f_{Y|X}(y|x_b; \hat{\theta}_{y|x})$ 로부터 무작위로 추출하여 대체된다. 이때, 무작위 추출을 위해 사용되는 분포에서 미지의 모수 $\theta_{Z|X}$ 와 $\theta_{Y|X}$ 는 데이터 B와 A를 이용하여 추정된다. 대표적으로 사용되는 모형 $f(\cdot)$ 으로는 다변량

정규분포와 회귀모형이 고려된다. 또한 Y 혹은 Z 가 이항변수인 경우에는 로지스틱 회귀모형이 사용되고 예측된 성공 확률이 0.5 이상인 경우에는 1을 그리고 0.5 이하인 경우에는 0을 부여하는 방안이 사용된다. 경험적 정보가 있는 경우에는 0.5 대신 다른 컷오프 값이 적용될 수 있다.

각각의 조건부 분포로부터 무작위로 생성된 값이 대체되기 때문에 모수적 모형에 대한 가정하에서는 결합된 연계 데이터를 이용한 주변 및 결합 확률분포에 대한 추론은 조건부 평균 매칭과는 달리 통계적으로 타당하다.

나. 비모수적 마이크로데이터 연계 방안

비모수적 방법을 이용한 마이크로데이터 연계에서는 변수(X, Y, Z)에 대한 특정한 모수적 분포를 가정하지 않고 데이터 A와 데이터 B를 연계하여 결합 마이크로데이터를 생성한다.(Okner, 1972).

비모수적 마이크로데이터 연계 방안은 크게 두 가지로 구분할 수 있다. 하나는 조건부 독립의 가정하에서 (X, Y, Z)의 분포를 비모수적으로 추정하고, 추정된 분포로부터 무작위로 각 결측을 대체하기 위한 값을 추출하는 방안으로, 모수적 방안과 일부 유사한 부분이 있다. 그러나 이러한 방안은 비모수적 분포 추정의 효율성 그리고 무작위 추출 방안의 선택 등에 의존하게 되어 마이크로데이터 연계를 위한 방안으로는 크게 선호되지 않는다.

다른 하나인 마이크로데이터 연계방안은 제공 파일의 개체 중 기준 파일의 각 개체와 주어진 기준하에서 유사한 개체를 선택하고 선택된 개체의 값으로 기준 파일에서 측정되지 않은 값을 대체하는 방안이다. 이 방안 역시 연계를 위해 사용되는 공통 변수 벡터 X 가 주어졌을 때, 두 변수

벡터 Y 와 Z 의 조건부 독립의 가정하에서 이루어진다.

여기에서는 개체 간 유사성을 이용한 여러 가지 마이크로데이터 연계 방안 중 실제 많이 사용되는 핫덱 대체(hot-deck imputation) 방안을 살펴본다. 핫덱 대체는 특별한 모수 분포의 가정 없이 X 를 바탕으로 하여 제공 파일과 기준 파일에 속한 개체들의 거리를 정의하고 이를 바탕으로 하여 무작위로 제공 파일의 개체를 선택하여 기준 파일의 값을 대체하는 방안이다. 랜덤 핫덱은 그 사용 방식에 따라 크게 랜덤 핫덱(random hot deck), 순위 핫덱(rank hot deck), 거리 핫덱(distance hot deck)의 세 가지로 구별할 수 있으며 이에 대한 간단한 설명은 다음과 같다 (Singh et al., 1993).

1) 랜덤 핫덱(Random hot deck)

랜덤 핫덱 방안은 제공 파일로부터 무작위로 개체를 선택하여 선택된 개체의 값으로 기준 파일의 결측을 대체하는 방안으로, 일반적으로 유사한 개체로 대체하기 위해서 제공 파일 및 기준 파일을 X 변수 기준으로 그룹화한 후 수행한다. 그룹을 나누기 위하여 흔히 사용되는 변수로는 성, 연령, 지역과 같은 기본적인 사회·인구학적 변인과 더불어, 개인 데이터의 경우에는 각 개인의 특성을 나타내는 소득과 교육수준과 같은 변수가 고려된다.

조건부 독립의 가정하에서 연계된 데이터를 이용한 각 변수의 주변 분포 및 결합 분포의 추정치는 통계적으로 타당하다. 즉, 랜덤 핫덱을 통해 연계된 데이터를 이용하여 (Y, Z) 의 모평균과 모분산과 같은 모수를 추정할 경우에 추정량의 비편향성을 포함한 기본적인 통계적 성질이 유지된다. 다만 사용되는 X 변수 중 연속형 변수가 있는 경우에는 데이터의 연

제 전에 이의 범주화가 필요하며 데이터 연계의 품질이 연속형 변수의 범주화 방안에 영향을 받을 수 있다.

랜덤 핫덱의 타당성은 조건부 독립의 가정이 성립하는 그룹의 구성 여부에 의하여 결정된다. 정확한 (X, Y, Z) 의 분포를 모르는 상황에서 조건부 독립성의 가정을 만족하는 그룹을 구성하기 위해서는 가능한 한 X 가 포함하고 있는 모든 정보를 활용하는 것이 바람직하며, 이에 따라서 적절한 크기의 그룹 내 개체 수를 유지할 수 있는 범위에서 그룹의 수를 최대화하는 것이 바람직하다.

2) 순위 핫덱(Rank hot deck)

순위 핫덱은 제공 파일과 기준 파일의 개체에 X 변수 기준으로 각각 순위를 부여하고 이를 이용하여 마이크로데이터를 연계하는 방안이다. 순위를 부여하기 위해서는 일반적으로 일변량 X 변수가 사용되며 공통변수 X 의 수가 많은 경우에는 이를 이용하여 순위를 결정하기 위한 생성변수를 우선 작성해야 한다. 가장 손쉬운 방안은 다음의 X 변수의 선형결합을 통해 생성변수 X_{new} 를 작성하는 것이다. 여기서 계수 c_i 는 각 변수의 상대적 중요도를 나타내는 값으로 $\sum_{i=1}^p c_i = 1$ 의 조건을 만족하도록 선택할 수 있다.

$$X_{new} = c_1 X_1 + c_2 X_2 + \cdots + c_p X_p$$

제공 파일의 크기가 기준 파일보다 크거나 같은 경우에는 부여된 순위를 직접적으로 이용하여 동일 순위 혹은 근접 순위를 갖는 제공 파일 내 개체의 값으로 기준 파일에 존재하는 결측을 대체할 수 있다. 또한 제공

파일과 기준 파일의 크기에 관계없이 경험 확률분포를 이용하여 결측을 대체할 수 있다. 이를 위해서는 우선 기준 파일과 제공 파일에서 X 의 경험 누적분포함수를 다음과 같이 추정한다.

$$\text{제공 파일: } \hat{F}_X^A(x) = \frac{1}{n_A} \sum_{a=1}^{n_A} I(x_a \leq x)$$

$$\text{기준 파일: } \hat{F}_X^B(x) = \frac{1}{n_B} \sum_{b=1}^{n_B} I(x_b \leq x)$$

여기서 n_A 와 n_B 는 각각 제공 파일과 기준 파일의 크기, 즉 개체 수를 나타낸다. 기준 파일에서 $X = x_b$ 값을 갖는 개체를 연계하기 위해서는 제공 파일에서 x_b 와 가장 유사한 개체, 즉 아래의 식을 만족하는 $X = x_a^*$ 를 갖는 개체를 연계를 위해 선택하게 된다.

$$\left| \hat{F}_X^A(x_a^*) - \hat{F}_X^B(x_b) \right| = \min_{1 \leq a \leq n_A} \left| \hat{F}_X^A(x_a) - \hat{F}_X^B(x_b) \right|$$

순위 핫덱을 통한 마이크로데이터 연계는 랜덤 핫덱과 더불어 상대적으로 작업이 쉬운 방안이나 공통변수인 X 의 수가 많은 경우에는 각 개체에 순위를 부여하거나 경험 누적분포함수를 추정하기 위한 추가적인 작업이 필요하며 그 방안에 따라 연계 결과가 달라질 수 있는 위험이 있다. 공통 변수들의 선형결합을 통해 일변량 생성 변수를 작성하고 이를 이용하여 순위 핫덱을 통한 연계 방안을 수행하는 간단한 대안을 고려할 수 있으나 이 역시 선형결합을 위한 계수 선택에 영향을 받을 수 있어 주의가 필요하다.

3) 거리 핫덱(Distance hot deck)

거리 핫덱은 통상적으로 많이 사용되는 마이크로데이터 연계 방안으로, 공통변수를 이용한 개체 간의 거리를 측정하여 기준 파일의 개체와 가장 근접한 제공 파일의 개체를 연계하는 방안이다(Okner, 1972; Rodgers, 1984).

즉, 두 개체의 거리를 나타내는 함수 $d(i,j)$ 가 거리함수가 갖추어야 하는 기본 성질을 만족한다고 가정한다. 일반적으로 거리함수가 갖추어야 하는 조건은 다음과 같다.

$d(i,j)$ is nonnegative, differentiable strictly convex function

$d(i,j) = 0$ if and only if $i = j$

$d'(i,j)$ is continuous

거리 핫덱을 이용한 마이크로데이터 연계는 주어진 거리함수를 이용하여 공통 변수 $\mathbf{X} = \mathbf{x}_b$ 를 갖는 기준 파일에 속한 개체의 $\mathbf{Y}$ 값을 대체하기 위하여 제공 파일에 속한 n_A 의 개체 중에서 $d(\mathbf{x}_a, \mathbf{x}_b)$ 를 최소화하는 $\mathbf{X} = \mathbf{x}_a^*$ 값을 갖는 개체를 선택하여 이의 $\mathbf{Y}$ 값을 대체하는 방안이다. 따라서 거리 핫덱 방안의 결과는 사용되는 거리함수에 의하여 결정된다. 일반적으로 사용되는 거리함수들은 다음과 같다.

■ 맨해튼(Manhattan) 거리 함수

$$d(i,j) = \sum_{k=1}^n |x_{ik} - x_{jk}|$$

■ 유클리디안(Euclidean) 거리 함수

$$d(i, j) = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2}$$

■ 마할라노비스(Mahalanobis) 거리 함수

$$d(i, j) = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{S}^{-1} (\mathbf{x}_i - \mathbf{x}_j)}.$$

여기서 $\mathbf{S}$ 는 분산 공분산 행렬의 추정량을 나타낸다.

■ 정확(Exact) 거리 함수

$$d(i, j) = \sum_{k=1}^n D_{(ij)k}, \quad D_{(ij)k} = \begin{cases} 1, & x_{ik} = x_{jk} \\ 0, & x_{ik} \neq x_{jk} \end{cases}$$

■ 비유사성 지수를 이용한 Gower 거리 함수

$$\text{dissimilarity } d(i, j) = \frac{\sum_{k=1}^n \delta_{ijk} d_{ijk}}{\sum_{k=1}^n \delta_{ijk}}.$$

여기서 δ_{ijk} 는

$$\delta_{ijk} = \begin{cases} 0, & \text{if } x_{ik} = \text{Missing} \text{ or } x_{jk} = \text{Missing}, \\ 0, & \text{if } x \text{ is binary variable and } x_{ik} = x_{jk} = 0 \text{ or } x_{ik} \neq x_{jk}, \\ 1, & \text{otherwise.} \end{cases}$$

이며 d_{ijk} 는 다음 표의 정의와 같다.

〈표 5-1〉 변수속성에 따른 거리함수

변수 속성	거리함수
범주형 변수 (nominal scale)	$d_{ijk} = \begin{cases} 0, & \text{if } x_{ik} = x_{jk}, \\ 1, & \text{otherwise.} \end{cases}$
연속형 변수 (interval scale)	$d_{ijk} = \frac{ x_{ik} - x_{jk} }{R_k},$ R_k는 k 번째 변수의 범위
순서형 변수 (ordinal scale)	$z_{ik} = \frac{(r_{ik} - 1)}{\max(r_{ik}) - 1},$ r_{ik}는 k 번째 변수기준 관측치 i의 상대적 위치. 새로운 변수 z_{ik}를 연속형으로 처리하여 거리를 정의함.

■ Minimax 거리함수

$$d(i, j) = \max(|x_{i1} - x_{j1}|, |x_{i2} - x_{j2}|, \dots, |x_{in} - x_{jn}|)$$

공통변수 X 의 단위가 유사한 경우에는 맨해튼 거리함수나 유클리디안 거리함수를 사용할 수 있으나 X 의 단위가 상이하여 이들의 분산이 유의하게 다르거나 공통 변수들 간의 유의한 상관관계가 있는 경우에는 이를 반영한 마할라노비스 거리함수를 사용하는 것이 바람직하다. 또한 공통 변수들 중 범주형 변수나 순서형 변수가 일부 포함된 경우에는 각 변수의 특성을 반영한 거리함수인 Gower 거리함수를 사용하는 것이 바람직하다.

다. 혼합 마이크로데이터 연계 방안

혼합 마이크로데이터 연계 방안은 앞에서 설명한 두 방안인 모수적 방안과 비모수적 방안을 단계적으로 적용하는 것으로 각 방안의 장점을 살리고 단점을 극복하기 위한 방안으로 흔히 고려된다. 실제 혼합 연계 방안의 적용을 위해 첫 번째 단계에서 사용되는 모수적 방안은 매크로 연계 방안으로 이해할 수 있다. 즉, 조건부 독립의 가정하에서 변수 (X, Y, Z) 의 결합분포를 가정하고 이의 모수를 추정하기 위하여 모수적 연계방안이 적용된다. 이후 모수적 방안을 통해 추정된 분포를 이용하여 비모수적 마이크로데이터 연계 방안을 사용하여 마이크로 연계 데이터가 생성된다.

예를 들어, (X, Y, Z) 의 분포로 다변량 정규분포를 가정하고 조건부 독립성의 가정하에서 다변량 정규분포의 모수인 모평균 벡터와 모분산-공분산 행렬을 추정한 후 제공 파일과 기준 파일에 존재하는 결측값을, 추정된 분포로부터 무작위로 추출, 대체하여 데이터 A_{imp} 와 B_{imp} 를 구성한다. 즉, A_{imp} 는 추정된 다변량 정규분포로부터 무작위로 생성된 Z 값으로 대체되어 생성되며 B_{imp} 는 추정된 다변량 정규분포로부터 무작위로 생성된 Y 값으로 대체되어 생성된다.

두 번째 단계에서는 대체가 이루어진 기준 파일과 제공 파일에 속한 개체들의 속성을 이용하여 랜덤 핫덱, 순위 핫덱 혹은 거리 핫덱을 수행하여 최종 마이크로 연계 데이터를 생성할 수 있다. 이때 데이터 A_{imp} 와 B_{imp} 에는 첫 번째 단계에서 대체과정을 통해 (X, Y, Z) 의 값이 모두 존재하기 때문에 여러 가지 방안들을 선택적으로 고려할 수 있다. 예를 들어, 세 변수 모두를 이용한 거리 핫덱 방안을 고려하거나 기준 파일에 포

합되지 않았던 Y 의 값만을 이용한 거리 핫텍 방안을 고려할 수 있다.

혼합 연계 방안은 모수적 모형이 비모수적 모형보다 더 보수적인 점과 비모수 방법이 모형 오류에 대해 더 강건한 점을 감안하여 각 방안의 단점을 보완하기 위하여 사용된다. 또한 비모수 방법의 핫텍 방법은 모형을 통해 추정된 예측 값을 대체하는 대신 제공 파일로부터 관측된 값을 대체함으로 모수적 방안에서 발생할 수 있는 비현실적인 값의 대체를 막을 수 있다.

제2절 데이터 연계 평가 및 쟁점

공공분야에서 이용자들에게 신뢰할 수 있는 데이터를 제공하기 위해서는 연계된 통합 데이터에 대한 평가가 필수적이다. 여기에서는 데이터 연계 결과에 대한 평가 기준과 고려 사항을 알아보고, 통합 데이터에서 중요한 관심사인 가중치 산정 방법에 대해 살펴보고자 한다. 그리고 기존 데이터 마이크로 연계 방안들에 대해 이론적으로 고찰하고, 기존 방안들의 한계를 간단한 모의실험을 통해 살펴서 이를 극복하기 위해 필요한 연구 방향을 제시하려 한다.

1. 데이터 연계 결과의 평가

앞서 소개된 여러 마이크로데이터 연계 방안들을 통해 생성된 결합 마이크로데이터의 품질은 기본적으로 조건부 독립성의 가정과 각 방안의 수행을 위해 고려되는 모형에 대한 가정의 타당성에 의하여 결정된다. 따라서 이러한 점을 감안하여 연계를 통해 생성된 데이터는 사용 전에 데이

터의 품질에 대한 적절한 검토가 선행되어야 할 것이다. 마이크로데이터 연계를 통해 생성된 데이터 및 이를 이용한 분석 결과의 평가를 위해서는 다음의 사항들을 고려해야 한다(D'Orazio et al., 2006).

1. 변수 (X, Y, Z)을 위한 모형 혹은 가정의 타당성
2. 결합 확률함수 $f(x, y, z)$ 의 추정량의 정확성
3. 사용된 마이크로데이터 연계 방안의 적절성
4. 결합 마이크로데이터를 이용한 추론의 타당성

여기서 항목 1, 2와 3은 매크로 연계 및 마이크로 연계 방안 모두에서 평가를 위해 고려되어야 하며 항목 3과 4는 특별히 마이크로데이터 연계 방안의 평가를 위해서 고려되어야 한다.

먼저 (X, Y, Z)의 모형에 대한 타당성 평가는 매우 어려운 부분이다. 이는 연계를 위한 데이터들이 모형을 평가하기 위한 최소한의 정보도 포함하고 있지 않기 때문이다. 즉, 두 데이터에 포함된 정보로는 (X, Y)와 (X, Z)의 분포에 대한 검증은 가능하나 (X, Y, Z)의 분포에 대한 가정을 위한 검증은 가능하지 않기 때문이다. 특별히 모형에 대한 가정 중 가장 중요한 역할을 하는 조건부 독립성 가정에 대한 평가는 반드시 이루어져야 하나 각 방안별 검증 방법에 대한 연구는 매우 취약하다. 이는 무응답 처리를 위한 통계적 방안들이 대부분 Missing at random(MAR) 가정하에서 제안되고 있으나 이에 대한 검증이 이루어지지 못하는 것과 동일하다.

추정량의 정확성에 대한 평가는 주어진 모형 및 가정하에서 전통적인 평가 방안들을 통해 상대적으로 용이하게 이루어질 수 있다. 대표적인 추정량의 평가 도구로는 분산 및 평균제곱오차와 같은 추정량의 변동성을

나타내는 통계와 대표본 성질(large sample property)을 나타내는 일치성 및 근사 정규성(asymptotic normality) 등을 고려할 수 있다.

마이크로데이터 연계 방안에 대한 평가는 우선적으로 연계된 데이터의 대표성 측면에서 일반적으로 이루어진다. 연계가 이루어진 데이터의 대표성을 평가하기 위하여 고려해야 하는 내용들은 다음과 같다.

- 연계된 데이터의 값과 참값의 일치 여부
- 상관관계를 포함한 변수들의 주변 및 결합 분포의 타당성

위에 언급된 내용을 실제로 검증하고 판단하는 것은 쉽지 않다. 연계된 데이터를 이용한 (X, Y) 와 (X, Z) 의 분포에 대한 검증은 연계 이전 자료와의 비교를 통해 가능하나 (Y, Z) 및 (X, Y, Z) 에 대한 검증은 필요한 정보가 존재하지 않아 매우 어려운 작업이다. 다만, 혼합 연계 방안이 적용된 경우에는 참값 (Y, Z) 와 모수적 방법을 통해 생성된 (Y_{imp}, Z_{imp}) 의 값을 비교하여 연계 과정에서 발생하는 잡음(noise)을 추정하고 이를 근거로 한 연계 방안의 타당성 검증을 일부 할 수 있는데, 이에 관한 연구 역시 매우 부족하여 추가적으로 연구할 필요가 있다.

2. 마이크로 연계 데이터 분석을 위한 가중치

조건부 독립의 가정하에서 독립적인 두 조사자료를 통해 작성된 데이터 파일 A와 B의 연계를 통해 생성한 마이크로데이터의 분석을 위해서는 가중치의 작성이 필수적이다. 이는 조사자료 데이터의 수집 목적이 유한 모집단의 분석에 있으며, 이를 위해서는 표본의 각 개체가 대표하는 모집단 개체수를 나타내는 가중치가 적용되어야 하기 때문이다. 이러한 가중

치의 사용은 추정량의 (근사) 비편향성을 확보하기 위한 최소한의 조건이다. 여기에서는 연계된 마이크로데이터의 분석을 위해 필요한 가중치에 대한 Rubin(1986)의 연구 결과를 바탕으로 하여 가중치 산출 방안을 설명한다.

Rubin(1986)은 마이크로데이터 연계를 통해 생성된 두 파일의 합집합인 $A \cup B$ 를 고려하였다. 만약 파일 A 혹은 B만을 이용한 분석, 즉 하나의 파일을 기준 파일로 그리고 또 다른 파일을 제공 파일로 정의할 경우에는 기준 파일에 존재하는 가중치를 이용하여 무한 모집단에 대한 분석을 실시할 수 있기 때문이다. 즉, 앞에서 고려한 여러 연계 방안들 중 반드시 제공 파일과 기준 파일이 구분되어야 하는 경우에는 특별히 가중치 문제를 고려하지 않아도 된다. 그러나 기준(recipient)과 제공(donor)의 역할이 교환되어 수행되고 이를 통해 변수의 수뿐 아니라 표본의 수 역시 증대하는 연계 방안이 고려된 경우에는 이의 분석을 위한 가중치의 산출은 반드시 필요하다.

Rubin(1986)은 파일 A 혹은 파일 B에 속한 개체들의 표본포함확률을 이용하여 연계된 마이크로데이터 분석을 위한 표본 가중치 w_a^{AB} 와 w_b^{AB} 를 산출하는 방안을 제안하였다. 새로운 가중치 산출 방안은 연계를 통해 생성된 표본 $A \cup B$ 의 가중치를 계산하기 위해 개별 단위들의 두 표본 중 적어도 하나의 표본에 포함될 확률 $\pi_s^{A \cup B} (= \pi_s^A + \pi_s^B)$ 를 이용하여 다음과 같이 산출된다.

- 조사 A에서의 개별 단위에 대한 새로운 가중치

$$w_a^{AB} = \frac{1}{\pi_a^{A \cup B}} = \frac{1}{1/w_a^A + 1/w_a^B}$$

- 조사 B에서의 개별 단위에 대한 새로운 가중치

$$w_b^{AB} = \frac{1}{\pi_b^{A \cup B}} = \frac{1}{1/w_b^A + 1/w_b^B}$$

여기서 π_s^A 와 π_s^B 는 각각 조사자료 A를 위한 표본과 조사자료 B를 위한 표본에 s 에 속한 개체가 포함될 확률을 나타낸다. 즉, 조사자료가 독립적으로 수행되었으며 동일한 유한 모집단으로부터 표본이 추출된 것으로 가정할 경우에 한 개체가 $A \cup B$ 에 속할 확률은 두 조사자료를 위한 표본 중 하나의 표본에 포함될 확률과 동일하며 이는 두 조사자료를 위해 계산된 개체 포함확률의 합으로 표현된다는 것이다.

언급한 과정을 통해 생성된 가중치는 기본적으로 각 개체가 서로 다른 두 조사를 위한 표본에 포함될 확률에 근거한 설계 가중치(sampling weight)로 이해할 수 있으며, 이중 표집틀(dual frame) 조사자료의 자료 분석을 위해 일차적으로 생성되는 설계 가중치와 동일한 방식을 통해 유도되었다. 두 조사자료의 분석 대상 모집단이 동일하다는 가정하에서 분석을 위한 최종 가중치의 산출을 위해 제안된 설계 가중치를 조정할 수 있다. 유한 모집단에 대한 보조정보를 이용한 가중치 조정 방안으로는 사후층화, 레이킹, 회귀 방안 등이 흔히 고려된다(Déville and Särndal, 1992).

3. 마이크로데이터 연계의 한계

앞에서 언급한 여러 가지 연계방안들의 사용과 이를 통해 생성된 데이터의 사용을 위해서는 기본적으로 조건부 독립에 대한 가정이 만족되어야 한다. 특별히 데이터 연계의 목적이 두 데이터에 함께 포함되지 않은

(Y, Z) 및 (X, Y, Z) 의 분포에 대한 분석에 있는 것을 감안할 때 조건부 독립의 가정은 매우 강력한 가정이며, 연계된 데이터의 분석방안에 따라서는 마이크로데이터 연계 자체의 필요성에 대한 재검토를 요구하는 사항이 될 수 있다.

예를 들어, 연계된 마이크로데이터를 이용하여 다음과 같은 다중회귀 모형(multiple regression model)을 적합하는 경우를 고려하자.

$$y = \beta_0 1 + x\beta_x + z\beta_z + \epsilon$$

$$E(\epsilon) = 0, \quad Var(\epsilon) = \sigma^2 I.$$

만약 조건부 독립의 가정, $Cov(Y, Z | X) = 0$ 이 만족되면 위 다중회귀 모형에서 x 가 주어졌을 때 y 와 z 의 선형관계의 정도를 나타내는 회귀 계수 $\beta_z = 0$ 이며, 따라서 마이크로데이터 연계의 목적인 다중회귀분석은 그 의미를 상실하게 된다. 즉, 연계된 자료를 이용한 인과관계(causal relationship) 혹은 연관관계(association) 분석에서 두 데이터에 각각 존재하는 두 변수 y 와 z 의 x 가 주어졌을 때 조건부 연관관계는 조건부 독립성 가정에 의하여 존재하지 않게 된다. 즉, 조건부 독립성의 가정은 연계된 자료의 분석에서 중요한 제약 조건이 되는 문제가 발생한다.

여기에서는 실제 조건부 독립의 가정이 만족되지 않음에도 불구하고 연계가 이루어진 경우에 조건부 독립의 가정이 마이크로데이터 연계 결과에 미치는 영향을 살펴보기 위하여 수행한 간단한 모의실험 결과를 살펴보고자 한다.

모의실험을 위해서 조건부 독립 가정이 만족하는 경우와 그렇지 않은 경우를 고려하여 두 종류의 데이터를 생성하였다. 이를 위해서 Y 와 Z 모두 X 의 선형함수에 오차를 더해 생성하되 오차항의 독립에 대한 가정을

달리하여 Y 와 Z 를 생성함으로 조건부 독립이 만족하는 경우와 그렇지 않은 경우의 데이터를 생성하였다. 각 데이터를 생성하기 위하여 사용한 구체적인 모형은 각각 다음과 같다.

■ 조건부 독립을 만족하는 데이터 생성 모형

$$Y = 3X + \epsilon_1,$$

$$Z = 2X + \epsilon_2,$$

$$X \sim N(6, 1), \quad X \perp \epsilon_1$$

$$\epsilon_1 \sim N(0, 1), \quad \epsilon_2 \sim N(0, 1), \quad \epsilon_1 \perp \epsilon_2$$

■ 조건부 독립을 만족하지 않는 데이터 생성 모형

$$Y = 2X + \epsilon_1$$

$$Z = 2X + \epsilon_2,$$

$$X \sim N(6, 1), \quad \epsilon_1 = 2\epsilon_2$$

$$\epsilon_1 \sim N(0, 1), \quad \epsilon_2 \sim N(0, 1)$$

데이터 연계 방안으로는 X 를 이용하여 크기가 10,000인 두 데이터를 각각 50개의 그룹으로 나누고 각 그룹정보를 이용한 랜덤 핫덱 방안과 조건부 평균 핫덱을 고려하였다. 즉, 각 그룹 내에서 랜덤 핫덱과 평균 핫덱을 적용하여 Y 와 Z 값을 대체하여 연계된 데이터를 작성하였다. 연계된 데이터를 통해 계산된 조건부 상관계수에 대한 결과는 다음의 표와 같다.

〈표 5-2〉 조건부 상관계수 결과

$Corr(Y, Z X)$	참값	랜덤 핫텍 연계	평균 핫텍 연계
Conditionally Independent	-0.0037	-0.0037	0.0065
Conditionally Dependent	0.9999	0.0014	0.8950

〈표 5-2〉의 값들은 X 가 주어졌을 때 계산된 Y 와 Z 의 조건부 상관계수를 나타낸다. 첫 행은 조건부 독립의 조건이 만족하는 데이터의 결과를, 둘째 행은 그렇지 않은 경우의 결과를 나타내고 있다. 참값 열은 실제 조건부 상관계수로, 모의실험의 경우에는 모든 변수의 참값을 알기 때문에 계산할 수 있다. 두 연계 방안을 통해 생성된 마이크로 연계 데이터를 이용하여 계산된 조건부 상관계수는 각각 랜덤 핫텍 연계 행과 평균 핫텍 연계 행에 주어져 있다.

조건부 독립성 가정이 만족된 경우에는 두 연계 방안 모두에서 조건부 상관계수에 대한 타당한 결과가 산출되었음을 확인할 수 있다. 그러나 조건부 독립의 가정이 만족되지 않는 경우에는 두 방안이 매우 상이한 결과를 제공하고 있다. 특히 랜덤 핫텍 방안은 조건부 독립의 가정으로 인하여 실제 조건부 상관계수가 거의 1임에도 불구하고 연계된 데이터를 이용하여 산출된 결과는 조건부 상관계수가 0.0014로 나타나는 것을 확인할 수 있다. 상대적으로 평균 핫텍은 비록 과소 추정을 하고는 있으나 랜덤 핫텍 방안보다 타당한 결과를 제공하고 있음을 확인할 수 있다. 다만, 본 절에서 고려한 모의실험은 매우 제한적이며 이러한 결과를 일반화하기 위해서는 추가적인 이론연구 및 실증연구가 심도 있게 진행되어야 할 것이다.

현재까지의 마이크로데이터 연계 방안에 대한 연구는 주로 주어진 조

건부 독립의 가정하에서 서로 다른 연계 방안들의 비교와 실제 적용 결과 분석 및 활용 가능성 부분에 국한되어 왔다. 그러나 앞서 살펴본 바와 같이 조건부 독립성의 가정이 만족될 경우에는 연계 데이터 생성의 목적인 다양한 연관관계 분석이 제한적이며, 또한 조건부 독립의 조건이 만족하지 않는 경우에는 연관관계 분석에 대한 결과의 통계적 타당성이 심각하게 훼손될 수 있다. 이러한 사실을 고려하여 추후에는 조건부 독립성 가정의 만족 여부와 상관없이 (Y, Z) 및 (X, Y, Z) 의 분포 내지 공분산 구조를 일정 수준 이상으로 추정할 수 있는 마이크로데이터 연계방안에 대한 연구가 이루어져야 할 것이다. 또한 조건부 독립의 가정이 만족되는 경우에도 마이크로 연계 과정에서 발생하는 추가적인 분산을 추정할 수 있는 방안에 대한 연구 역시 이루어져야 할 것이다.

제 6 장

결론 및 정책 제언

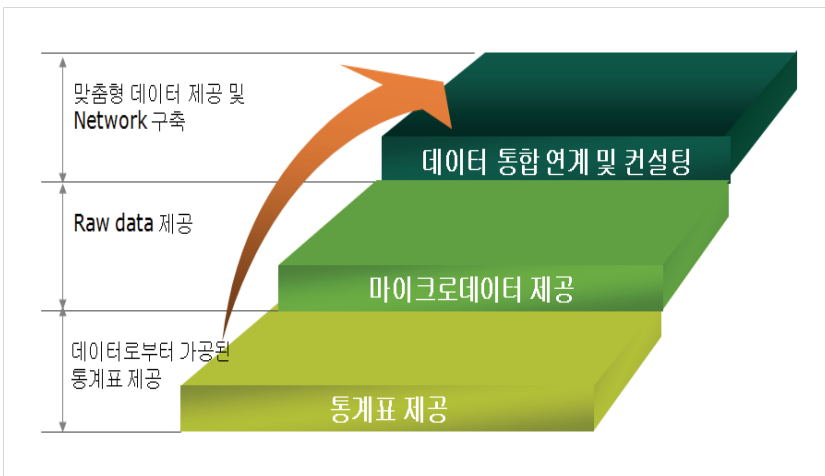
2016년 보건복지통계정보 통합 관리 및 운영 연구 사업에서는 2014년에 개편된 데이터 중심의 보건복지데이터포털을 각 항목별로 업데이트하고, 고도화하는 작업을 진행하였다. 보건복지데이터포털 이용자의 과거 행동 패턴 데이터를 바탕으로 하여 신규 이용자에게 필요한 정보를 제시하여 추천 자료를 제공하고자 추천 시스템을 개발하였으며, Data Bank의 Deposit data 섹션을 구축하여 향후에 원외 연구자들에게 ‘데이터 위탁관리 및 공유 서비스’를 제공하고자 하였다. 2015년에 구축한 인포그래픽스와 통계시각화는 올해에도 지속적으로 콘텐츠를 제공하고자, 연구원에서 생산한 2015년 전국 출산력 및 가족보건·복지 실태조사 결과를 중심으로 하여 작업하였다.

데이터에 대한 수요자의 요구와 관심이 높아지고 있는 이 시점에서, 보건복지데이터포털의 고도화 방향은 추천 시스템의 고도화와 마이크로데이터 연계 연구라고 할 수 있다. 추천 시스템은, 동일 저자의 다른 자료 추천, 이 자료와 같이 본 자료 추천 등의 기본적인 정보만을 제공해 주는 단계에서 나아가 이를 확장하여 이용자별로 실제 이용하는 페이지들에서 노출되는 주제와 관련된 추천 기능도 개발하여야 한다. 보건복지데이터포털에 적용할 수 있는 추천 시스템의 난이도를 조절하고 단계적으로 확장하여 이용자들에게 향상된 시스템으로 지속적인 서비스를 제공해야 할 것이다.

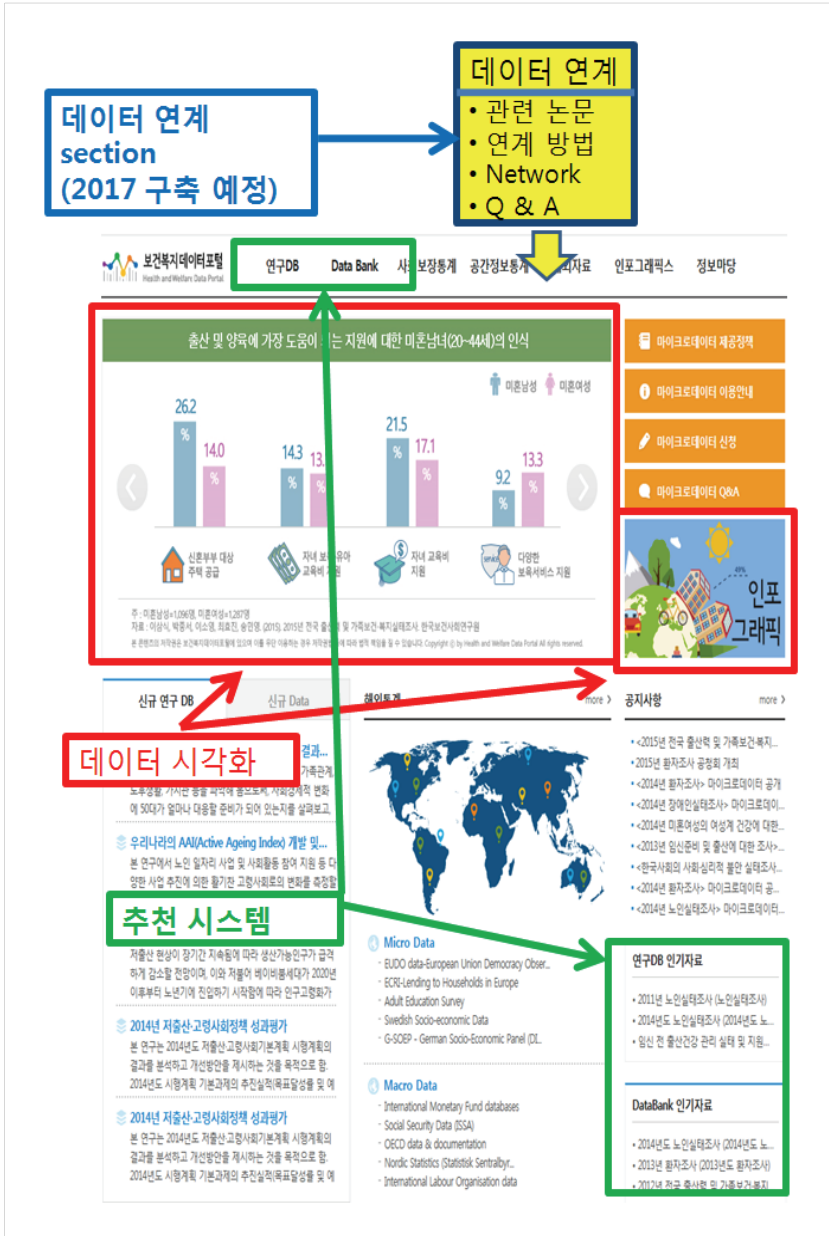
마이크로데이터 연계는 조사환경의 악화로 인해 어려워진 데이터 수집의 효율적인 통계 생산 대안으로 제시할 수 있다. 기존 포털은 데이터로부터 가공된 통계표를 제공하는 메타정보 제공 측면이 강했다면, 현재는

다양한 마이크로데이터를 제공하여 이용자가 직접 심층 분석 할 수 있는 환경을 구축한 데이터 제공 서비스가 중심이다. 향후 보건복지데이터포털의 방향성은 맞춤형 데이터 제공 및 네트워크 구축이라고 할 수 있다. 맞춤형 데이터 제공은 단일의 데이터 제공뿐만 아니라, 이용자가 원하는 여러 데이터와 변수를 선택하면 신뢰성 있는 통합데이터를 제공하는 것을 의미한다. 앞서 5장에서 언급한 것처럼, 통합데이터 제공에서 중요한 부분은 데이터 연계 결과에 대한 평가와 통합 데이터의 가중치 부여이다. 이를 위해서, 데이터 연계에 관심 있는 연구자들로 구성된 네트워크를 통해 마이크로데이터 통합 연계 활용사례 발표와 한계점을 논의할 수 있는 장을 마련하는 것이 필요하다. 보건복지데이터포털은 앞으로 마이크로데이터 연계 섹션(data linkage section)을 구축하여 마이크로데이터 통합 연계 연구와 관련된 논문 제공, 데이터 연계 방법 공유, 매뉴얼 개발, 질 의응답(Q&A) 등으로 데이터 연계 네트워크의 중심점이 되는 중요한 역할을 할 수 있다.

[그림 6-1] 보건복지데이터포털의 방향성



[그림 6-2] 보건복지데이터포털의 구성도 2



올해의 보건복지데이터포털 주요 연구내용과 방향성을 현재 포털 홈페이지에서 살펴보면 [그림6-2]와 같다. 기존의 보건복지데이터포털 서비스인 연구DB, Data Bank, 사회보장통계, 공간정보통계, 해외자료, 인포그래픽스 section은 업데이트 및 개편을 통해 유지하되, 포털 이용자의 축적된 정보로 연구DB 및 Data Bank section에 추천 시스템을 구축하여 이용자 맞춤형 정보를 제공하고자 하였다. 추천 시스템은 다양한 방법으로 고도화할 수 있기 때문에 4장의 이론적 검토를 바탕으로 하여 단계적으로 확장될 수 있도록 한다. 그리고 앞서 언급한 것처럼 데이터 연계에 대한 수요가 증가함에 따라, 2017년에 데이터 연계 section을 구축하여 6장의 이론적 내용 및 관심사 등을 연구자들이 공유할 수 있도록 구축할 예정이다.

기존의 포털시스템 서비스, 추천 시스템, 데이터 연계 section은 유기적인 연관성을 가질 수 있다. 기존의 포털시스템으로 축적된 정보를 이용하여 추천 시스템을 구축하고, 구축된 추천 시스템을 통해 인기가 높은 데이터를 파악할 수 있다. 또한, 추천 시스템으로 연관성을 분석하여 연관성 높은 연구과제 및 데이터를 연계할 수 있는 데이터 후보군으로 선정하여 통합데이터를 구축하고, 의미 있는 결과를 통계 시각화 및 인포그래픽으로 제공하도록 한다.

보건복지데이터포털에서 앞으로 연구적인 성격을 강화하고 연구자들이 소통할 수 있는 공간을 마련하여 다양한 정보를 제공한다면, 통계청의 국가통계포털(KOSIS), 사회과학자료원의 사이트와 차별화될 수 있을 것이다. 또한, 보건복지통계정보 서비스 환경 고도화 및 지속 가능한 시스템 운영을 위해, 운영진은 주기적으로 생산되는 자료의 업데이트와 신규 자료의 등록을 꾸준히 하여 이용자들에게 질 높은 통계서비스를 제공해야 할 것이다. 보건복지데이터포털 콘텐츠와 User Interface는 이용자들

의 만족도를 높일 수 있는 중요한 사항이므로 이 부분에 대해 끊임없이 고민하고 노력을 기울인다면, 보건복지데이터포털은 보건복지 분야에서 독보적인 존재로 자리매김할 수 있을 것이다.

참고문헌 <<

- 고은애. (2004). 통계적 매칭을 이용한 데이터 통합에 관한 연구. 석사학위논문, 동국대학교 대학원.
- 박창이, 김용대, 김진석, 송종우, 최호식. (2015). R을 이용한 데이터 마이닝, 개정판. 교우사.
- 안일호. (2003). 혼합형 데이터의 통계적 결합에 관한 연구. 석사학위논문, 고려대학교 대학원.
- 오미애, 이연희, 이상영, 오영희, 김문길, 윤강재, 진달래. (2013). 보건복지통계 정보 시스템 구축 및 운영. 한국보건사회연구원.
- 오미애, 최현수, 김용대, 이용희, 진재현. (2014). 보건복지통계정보 생산 및 활용 촉진을 위한 마이크로데이터 통합 연계 방안. 한국보건사회연구원.
- 오미애, 장영식, 최현수, 진재현, 천미경. (2015). 보건복지통계정보 통합 관리 및 운영. 한국보건사회연구원.
- 이병희. (2008). 저소득 노동시장의 실태와 동태적 변화. 동향과 전망, pp.204-230.
- 정성석, 김순영, 김현진. (2004). 데이터 보강을 위한 데이터 통합기법에 관한 연구. 응용통계연구, 제17권 3호, pp. 605-617.
- 통계개발원. (2013). 다양한 출처 자료 처리 및 통계 생산방안 연구.
- Adomavicius, G. and Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions, *IEEE Transactions on Knowledge and Data Engineering*, 17, pp.734-749.
- Agrawal, R. and Srikant, R. (1994). Fast algorithms for mining association rules, *Proceedings of the 20th International Conference on Very Large Data Bases (VLDB '94)*, Morgan Kaufmann, pp.487-499.

- Balabanovic, M. and Shoham, Y. (1997). Fab: content-based, collaborative recommendation, *Communications of the ACM*, 40, pp.66-72.
- Bayardo Jr, R. J. (1998). Efficiently mining long patterns from databases, *Proceedings of the 1998 ACM SIGMOD international conference on Management of data*, pp.85-93.
- Bishop, C. M. (1995). *Neural networks for pattern recognition*, Oxford university press.
- Bonett, D. G and Wright, T. A. (2000). Sample size requirements for Pearson, Kendall, and Spearman correlations, *Psychometrika*, 65, pp.23-28.
- Breese, J. S., Heckerman, D. and Kadie, C. M. (1998). Empirical analysis of predictive algorithms for collaborative filtering, *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence* (Madison, WI)(Gregory F. Cooper and Serafin Moral, eds.), Morgan Kaufmann, pp.43-52.
- Chakrabarti, S. (2002). *Mining the web: Discovering knowledge from hypertext data*, Science and Technology Books.
- Chen, Y.-H. and George, E..I. (1999). A Bayesian model for collaborative filtering, *Proceedings of Uncertainty 99: The Seventh International Workshop on Artificial Intelligence and Statistics* (Fort Lauderdale, FL).
- Cox, D. R. and Snell, E. J. (1989). *The Analysis of Binary Data*, 2nd Edition, London: Chapman & Hall.
- Deerwester, S., Dumais, S.T., Furnas, G. W. Landauer, T. K. and Harshman, R. (1990). Indexing by latent semantic analysis, *Journal of the American Society for Information Science*, 41, pp.391-407.

- Déville, J. C. and Särndal, C. E. (1992). Calibration Estimators in Surevy Sampling, *Journal of the American Statistical Association*, 87, pp.376-382.
- D'Orazio, Marcello, Di Zio, Marco and Scanu, Mauro. (2006). *Statistical Matching Theory and Practice*, Wiley.
- Friedman, J. H. (1997). On bias, Variance, 0/1-loss, and the curse-of-dimensionality, *Data Mining and Knowledge Discovery*, 1, pp.55-77.
- Fu, X., Budzik, J. and Hammond, K. J. (2000). Mining navigation history for recommendation, *Proceedings of the 5th International Conference on Intelligent User Interfaces (IUI '00)(New Orleans)*, ACM, pp.106-112.
- Grigorik, I. (2007). *SVD Recommendation system in ruby*, web blog, "<http://www.igvita.com/2007/01/15/svd-recommandation-system-in-ruby/>"
- Herlocker, J. L., Konstan, J. A. and Riedl, J. T. (2002). An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms, *Information retrieval*, 5, pp.287-310.
- Herlocker, J. L., Konstan, J. A., Terveen, L. G. and Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems, *ACM Transactions on Information Systems (TOIS)*, 22, pp.5-53.
- Ingram, D., O'Hare, J., Scheuren, F. and Turek, J. (2000). Statistical matching: a new validation case study, *Proceedings of the survey Research Methods Section, American Statistical Association*.
- Jannach, D., Zanker, M., Felfernig, A. and Friedrich, G. (2011). *An introduction: recommender systems*, Cambridge university press.
- Joachims, T. (1998). Text categorization with support vector machines: learning with many relevant features, *Proceedings of*

- the 10th European Conference on Machine Learning (ECML-98)(Chemnitz, Germany)(Claire Nedellec and Celine Rouveirol, eds.)*, no.1398, Springer Verlag, Heidelberg, Germany, pp.137-142.
- Kim, J. W., Lee, B. H., Shaw, M. J., Chang, H.-L. and Nelson, M. (2001). Application of decision tree induction techniques to personalized advertisements on internet storefronts, *International Journal of Electronic Commerce*, 5, pp.45-62.
- Koren, Y., Bell, R. and Volinsky, C. (2009). Matrix factorization techniques for recommender systems, *Computer*, 42, pp.30-37.
- Lemire, D. and Maclachlan, A. (2005). Slope one predictors for online rating-based collaborative filtering, *Proceedings of the 5th SIAM International Conference on Data Mining (SDM '05)(Newport Beach, CA)*, pp.471-480.
- Lin, W. (2000). *Association rule mining for collaborative recommender systems*, Master's thesis, Worcester Polytechnic Institute.
- Lin, W., Alvarez, S. A. and Ruiz, C. (2002). Efficient adaptive-support association rule mining for recommender systems, *Data Mining and Knowledge Discovery*, 6, pp.83-105.
- Manning, C. D., Raghavan, P. and Schutze, H. (2008). *Introduction to information retrieval*, Cambridge University Press.
- McCallum, A. and Nigam, K. (1998). A comparison of event models for naive bayes text classification, *In AAAI-98 Workshop on Learning for Text Categorization (Madison, WI)*.
- Miyahara, K. and Pazzani, M. J. (2000). Collaborative filtering with the simple bayesian classifier, *Pacific Rim International Conference on Artificial Intelligence (Melbourne, Australia)*, pp.679-689.
- Mobasher, B., Dai, H., Luo, T. and Nakagawa, M. (2001). Effective

- personalization based on association rule discovery from web usage data, *Proceedings of the 3rd International Workshop on Web Information and Data Management (WIDM '01)(Atlanta)*, ACM, pp.9-15.
- Moriarity, C. (2009). *Statistical Properties of Statistical Matching: Data Fusion Algorithm*. VDM Publishing.
- Nelder, J. and Wedderburn, R. (1972). Generalized Linear Models, *Journal of the Royal Statistical Society. Series A*, 135, pp.370-384.
- Okner, B. (1972). Constructing a new data base from existing microdata sets: the 1966 merge file. In *Annals of Economic and Social Measurement*, Volume 1, number 3, pp.325-362, NBER.
- Pazzani, M. and Billsus, D. (1997). Learning and revising user profiles: The identification of interesting web sites, *Machine Learning*, 27, pp.313-331.
- Pazzani, M. J. and Billsus, D. (2007). Content-based recommendation systems, The Adaptive Web (Peter Brusilovsky, Alfred Kobsa, and Wolfgang Nejdl, eds.), *Lecture Notes in Computer Science*, vol.4321, Springer, pp.325-341.
- Porter, M. F. (1980). An algorithm for suffix stripping, *Program*, 14, pp.130-137.
- Quinlan, J. R. (1993). *C4.5: programs for machine learning*, Morgan Kaufmann, San Francisco.
- Ricci, R. and Rokach, L. and Shapira, B. (2011). *Introduction to Recommender Systems Handbook*, Springer, pp.1-35.
- Rodgers, W.L. (1984). An Evaluation of Statistical Matching, *Journal of Business and Economic Statistics*, 2, pp.91-102.
- Rubin, D. B. (1986). Statistical Matching using file concatenation with

- adjusted weights and multiple imputations, *Journal of Business and Economic Statistics* 4, pp.87-94.
- Salton, G. and Buckley, C. (1988). Term-weighting approaches in automatic text retrieval, *Information Processing and Management*, 24, pp.513-523.
- Salton, G., Wong, A. and Yang, C. S. (1975). A vector space model for information retrieval, *Journal of the American Society for Information Science*, 18, pp.613-620.
- Sarwar, B., Karypis, G., Konstan, J. A. and Riedl, J. (2000a). Application of dimensionality reduction in recommender systems - a case study, *Proceedings of the ACM WebKDD Workshop (Boston)*.
- Sarwar, B., Karypis, G., Konstan, J. A. and Riedl, J. (2000b). Analysis of recommendation algorithms for e-commerce, *Proceedings of the 2nd ACM Conference on Electronic Commerce (EC '00)(Minneapolis, MN)*, ACM, pp.158-167.
- Shyu, M.-L., Haruechaiyasak, C., Chen, S.-C. and Zhao, N. (2005). Collaborative filtering by mining association rules from user access sequences, *Proceedings of the International Workshop on Challenges in Web Information Retrieval and Integration (WIRI '05)(Washington, DC)*, IEEE Computer Society, pp.128-135.
- Singh, A. C., Mantel, H., Kinack, M., & Rowe, G. (1993). Statistical matching: use of auxiliary information as an alternative to the conditional independence assumption. *Survey Methodology*, 19(1), pp.59-79.
- Su, X. and Khoshgoftaar, T. M. (2009). A Survey of Collaborative Filtering Techniques, *Advances in Artificial Intelligence*, Article ID 421425, p.19.

- Van Der Putten, Peter, Kok, Joost N., and Gupta, Amar. (2002). Why the information explosion can be bad for data mining, and how data fusion provides a way out, *Second SIAM International Conference on Data Mining*, Arlington, April, pp.11-13.
- Van Pelt, X. (2001). *The Fusion Factory: A Constrained Data Fusion Approach. Master of Science*. Thesis, Leiden Institute of Advanced Computer Science, The Netherlands.

포털명

보건복지부 통계포털

http://stat.mohw.go.kr

통계포털

인기 검색어 | 아동학대

보건복지 통계DB 찾기

통계DB

통계발간물

알림정보

고객참여

통계업무지원

통계국민제안 | 자료문의

고객참여

통계국민제안

통계국민제안

자료문의

통계국민제안

꼭 필요한 자료를 효과적으로 제공하도록 하였습니다.

통계국민제안 시 참고해주세요!

- 보건복지부 통계 서비스의 발전을 위한 좋은 의견을 기다립니다.
- 자료이용에 관련한 불만한 점이나 개선해야 할 사항에 대해서 구체적으로 기재해주시기 바랍니다.
- 특정 있는 계사를 위해 실명확인결과를 거치고 있으며, 육상, 비행, 광고성글, 허위사실유로 등의 글은 게시자에게 불이익이 발생할 수 있습니다.
- 동일내용의 반복게재, 게시물 도배 등 오피미자의 정상적인 운영을 저해하는 경우 글쓰기 권한이 제한될 수 있음을 알려드립니다.

제안사항

검색

번호	제안사항	작성일	작성자	답변상태
56	테스트	2016.07.12		진행중
55	경상남도 시, 군 치아우식증 유행률을 알고 싶습니다~	2016.06.14	서영경	완료
54	주택별 분류에 대한 물어 차이	2015.12.16	정영철	완료
53	맞춤형 기초생활수급자 현황	2015.11.27	박상훈	완료
52	의료기관내 종사자의 자상사고율에 대해 알고 싶습니다.	2015.11.27	이상협	완료
51	결핵병원에 대해 알고 싶습니다.	2015.11.05	박준명	완료
50	우리나라 빈곤가정 통계에 대해서 궁금한 점이 있습니다.	2015.11.03	이성욱	완료
49	우리나라 전체 가정을 통가가정의 비율이 궁금합니다.	2015.10.29	소은혜	완료
48	우리나라 기생충증 의 유행률	2015.10.19	김동호	완료
47	장애인 운전자 현황	2015.10.06	임광화	완료

전체 [56] Page 1/6

통계포털

인기 검색어 | 요양병원

보건복지 통계DB 찾기

통계DB

통계발간물

알림정보

고객참여

통계업무지원

보건복지 연보/백서 | 주택별 발간물 | 주요자료 100선 | 통계용어사전

통계발간물

보건복지 연보/백서

통계발간물

통계연보

백서

꼭 필요한 자료를 효과적으로 제공하도록 하였습니다.

본 사이트는 대한민국 저작권법을 준수합니다.

웹사이트에 게시된 자료는 저작권법에 의하여 보호받는 저작물로서, 자료를 등록한 부서의 허락을 받은 후 이용하시기 바랍니다.

제목

검색

번호	제목	출처	등록일	조회 수
67	2015 보건복지통계연보	정책통계담당관	2015.12.02	15769
66	2014 보건복지통계연보	정책통계담당관	2014.12.08	35620
65	2013 보건복지통계연보	정책통계담당관	2013.11.29	38547
64	2012 보건복지통계연보	정책통계담당관	2012.11.28	28044
63	2011 보건복지통계연보	정책통계담당관	2011.11.07	27210
62	2010 보건복지통계연보	정책통계담당관	2010.12.10	24537
61	2009년도 보건복지가족 통계연보	-	2009.11.11	41537
60	2008년도 보건복지가족 통계연보	-	2008.11.05	27799
59	2007년도 보건복지통계연보	-	2008.06.24	22136
58	2006년도 보건복지통계연보	-	2008.06.24	13157

전체 [67] Page 1/7

■ 서비스 제공

- 통계DB: 통계지표별로 정보를 제공하며 통계청의 e-나라지표와 국가 통계포털(KOSIS)에 링크하여 서비스하고 있음.
- 통계발간물: 보건복지연보 및 백서, 주제별 통계관련 발간물 정보를 게시하여 첨부파일로 내려받을 수 있도록 하며, 주요지표 100선에 대해서는 연도별 수치와 그래프를 한눈에 볼 수 있도록 제공함(단, 엑셀 등의 파일 형태로는 제공하지 않음).

■ 특이 사항

- 통계국민제안, 자료문의가 가능한 고객참여 메뉴가 있어, 이용자들이 현재는 생산되지 않은 보건복지 통계지표를 제안할 수 있고, 필요한 통계지표를 습득할 수 있는 정보 또는 경로를 제공함.
- 많이 본 통계자료가 메인페이지에 표시되는 것이 장점임.
- 메인화면의 '보건복지통계DB 찾기'를 클릭하면 보건복지주제별, 기념일별, 통계명에 따라 이용자가 쉽게 검색할 수 있음.

성인지동계

[홈](#)

통계DB
테마통계
정보
알림
소개

[검색](#)

성명등지수
주요지표
통계자료
도표로 보기
지도로 보기

» 테마통계 » 성명등지수

성명등지수
주요지표
통계자료
도표로 보기
지도로 보기

» 테마통계 » 성명등지수

성명등지수
주요지표
통계자료
도표로 보기
지도로 보기

» 테마통계 » 성명등지수

성명등지수
주요지표
통계자료
도표로 보기
지도로 보기

» 테마통계 » 성명등지수

성명등지수
주요지표
통계자료
도표로 보기
지도로 보기

» 테마통계 » 성명등지수

성명등지수
주요지표
통계자료
도표로 보기
지도로 보기

» 테마통계 » 성명등지수

성명등지수
주요지표
통계자료
도표로 보기
지도로 보기

» 테마통계 » 성명등지수

성명등지수
주요지표
통계자료
도표로 보기
지도로 보기

» 테마통계 » 성명등지수

성명등지수
주요지표
통계자료
도표로 보기
지도로 보기

» 테마통계 » 성명등지수

성명등지수
주요지표
통계자료
도표로 보기
지도로 보기

» 테마통계 » 성명등지수

성명등지수
주요지표
통계자료
도표로 보기
지도로 보기

» 테마통계 » 성명등지수

성명등지수
주요지표
통계자료
도표로 보기
지도로 보기

» 테마통계 » 성명등지수

성명등지수

포털명

한국여성정책연구원
성인지통계 정보시스템

http://gsis.kwdi.re.kr

성인지통계

주요지표

통계자료

도표도

지도도

도움말

우울증상 경험률

주요지표 > 건강상태

2005

2007

2008

2009

2010

2011

2012

2013

여성

남성

구분

2005

2007

2008

2009

2010

2011

2012

2013

여성

19.2

15.6

18.3

19.2

16.9

16.1

15.9

남성

11.5

8.1

10.1

10

7.9

10.1

6.9

자료안내

출처

원출처

보건복지부, '국민건강영양조사'

GIS 출처

주요지표 > 건강상태 > 우울증상 경험률(성별)

관련 사이트

www.mohw.go.kr

주기

1년

지표안내

개념

19세 이상(연령 표준화) 남녀의 우울증상 경험률을 제시

의의

우울증상의 경험률을 파악함으로써 정신건강 개선 관련 정책의 기초 자료로 활용

해석

2005년~2013년의 우울증상 경험률을 살펴보면, 여성의 우울증상 경험률이 남성에 비해 높음

우울증상 경험률은 2005년 여성 19.2%, 남성 11.5%이었으나 증감을 반복하다가 2013년 여성 13.7%, 남성 6.6%로 2005년에 비해 다소 감소함

여성의 우울증상 경험률이 남성에 비해 높아 우울증상 집단 확대 등 정신건강을 개선하기 위한 정책방안이 필요함

이용상 유의점

19세 이상 남녀 인구를 2005년 추계인구로 연령표준화함

용어설명

우울증상경험률: 최근 1년 동안 연속적으로 2주 이상 일상생활에 지장이 있을 정도로 슬프거나 절망감 등을 느낀 분율(19세 이상 중 우울증상경험자/19세 이상 인구) X 100

한국여성정책연구원 관련 연구

2015년 여성가족정책조사(조경선 외, 2015)

달별의 여성의 일·가정 양립 활동과 건강영향 연구: 근본사안을 중심으로(김동식 외, 2013)

한국 여성의 여성건강증진 정책 추진현황을 위한 한국 사회연구(김영택 외, 2013)

성인지적 관점에서의 청소년 건강실태와 정책과제(김동식 외, 2012)

100세 시대 여성 고령자의 건강의 수준과 정책방안(김영택 외, 2011)

성인지적 관점에서의 사회: 우울증 집단 확대방안 - 성인지적 분석 및 소요예산 중심으로(김영택 외, 2010)

성인지통계 관련 통계표

스트레스 정도(성/연령 그룹별) - 국민건강영양조사

스트레스 정도(성/연령 그룹별) - 국민건강영양조사

스트레스 정도(성/연령 그룹별) - 국민건강영양조사

스트레스 정도(성/연령 그룹별) - 국민건강영양조사

최종수정일

2016.03.09

포털명

한국여성정책연구원
성인지통계 정보시스템

http://gsis.kwdi.re.kr

성명불지수

주요자료

통계리포트

도표로 보기

지도로 보기

> 데이터통계 > 통계리포트

2016년 지역구 국회의원 선거의 여성 당선자 비율 10.3%로 역대 선거 중 가장 높아

비율 보기

통계표 다운로드

여성 후보자 비율

여성 당선자 비율

출처 : 중앙선거관리위원회, '선거관리 58, 1' (단위 : %)

2000

2004

2008

2012

2016

3.1

2.7

5.6

4.1

11.9

5.7

7.0

7.7

10.5

10.3

지역구 국회의원 후보자 및 당선자 여성비율 (2000-2016)

2000년 이후 국회의원 선거의 지역구 후보자와 당선자 여성비율을 살펴보면, 2016년 여성 당선자 비율은 10.3%로 역대 지역구 국회의원 선거 중 가장 높았습니다. 2016년 지역구 국회의원 선거의 여성 후보자 비율은 10.5%로 2008년 11.9%보다 다소 낮았습니다.

지역구 국회의원 선거의 여성 후보자 비율은 2000년 3.2%에서 2004년 5.6%, 2008년은 11.9%로 꾸준히 증가하다가 2012년 7.0%로 일시적으로 감소하였습니다. 지역구 국회의원 선거의 여성 후보자 비율은 2016년 10.5%로 증가하였으나 2008년 수준에 못 미치지 못하였습니다. 한편, 지역구 국회의원 선거의 여성 당선자 비율도 꾸준히 증가하고 있어 2000년 2.2%, 2008년 5.7%, 2012년 7.7%, 2016년 선거에서 처음으로 10%대에 진입하였습니다.

지역구 국회의원 선거의 여성 후보자와 당선자 비율이 꾸준히 증가하고 있지만 국민의 절반을 차지하는 여성의 정치적 대표성 확대 측면에서 여전히 매우 낮은 수준이라 할 수 있습니다.

※ '통계리포트'의 그래프 및 글을 인용하실 경우 한국여성정책연구원 성인지통계(GSIS)가 출처임을 꼭 밝혀주시기 바랍니다.

2016.06.02

여권글

이전글이 없습니다.

다음글

2015년 현재 달평군 인허가소장 구애, 여성이 남성보다 더 자주 이용해

목록

간행물

마이크로데이터

자료실

> 정보 > 마이크로데이터

2012 전국다문화가족실태조사 신청

▶ 신청자료 기본정보

신청하시는 자료가 대한 정보입니다.

기관

여성가족부

연구책임자

연구책임자

연락처

02-3156-7162

이메일

songgy@kwdi.re.kr

신청자료

☐ 5종일괄
☐ 다문화 가구
☐ 다문화 가구원
☐ 결혼이민자, 귀화자 중
☐ 결혼이민자, 귀화자 중의 배우자
☐ 한 주=24세 청소년 자녀

가격

0원

신청하시는 자료가 맞습니까? 예 ☐

▶ 신청자 정보

신청하시는 문의 정보를 입력해주세요.

신청자 성명

한미경

신청자 이메일

sunlimoon@khsa.re.kr

신청자 핸드폰

010

소속 기관

소속 부서

활동 목적

☒ 정책연구 보고서용
☐ 논문연구용
☐ 교육용
☐ 기타용

보고서 및 논문 게재

연구 목적

※ 제출받은 통계자료는 제출받은 목적 이외의 용도로 사용하거나 공문사용자로 가져갈 이외의 차용에 임의로 가공하거나 판매할 수 없습니다.

연구목적에 대해서 상세하고 정확하게 기술하여 주시기고, 이 자료를 사용하실 모든 문헌을 정확하게 작성하여 주십시오.

연구목적용 기재하지 않을 경우 승인여 여하를 수 있습니다.

주요 연구내용

※ 주요 연구내용을 기재하지 않을 경우 승인이 여하를 수 있습니다.

공문사용자

※ 자료를 사용할 모든 사용자 정보를 정확하게 작성하여 주십시오.

공문사용자 추가

순서	사용자명	소속기관	보안각서동의
1			☐
2			☐

포털명	한국여성정책연구원 성인지통계 정보시스템	http://gsis.kwdi.re.kr	
공지사항 DB 업데이트 보도자료 이벤트 문의사항 알림 > DB 업데이트			
성인지통계의 통계DB 업데이트 소식을 알려드립니다. 10개 년 월 주 ~ 년 월 주 검색			
번호	제목	기간	등록일
122	5월 4주 통계DB 업데이트	2016.05.23 ~ 2016.05.27	2016.05.26
121	5월 3주 통계DB 업데이트	2016.05.16 ~ 2016.05.20	2016.05.18
120	5월 2주 통계DB 업데이트	2016.05.09 ~ 2016.05.13	2016.05.12
119	5월 1주 통계DB 업데이트	2016.05.02 ~ 2016.05.06	2016.05.04
118	4월 4주 통계DB 업데이트	2016.04.25 ~ 2016.04.29	2016.04.28
117	4월 3주 통계DB 업데이트	2016.04.18 ~ 2016.04.22	2016.04.22
116	4월 2주 통계DB 업데이트	2016.04.11 ~ 2016.04.15	2016.04.15
115	4월 1주 통계 DB업데이트 휴간 안내	2016.04.04 ~ 2016.04.08	2016.04.11
114	3월 5주 통계DB 업데이트	2016.03.28 ~ 2016.04.01	2016.03.31
113	3월 4주 통계DB 업데이트	2016.03.21 ~ 2016.03.25	2016.03.28
<< < 1 2 3 4 5 6 7 8 9 10 > >>			
공지사항 DB 업데이트 보도자료 이벤트 문의사항 알림 > 보도자료			
각 기관에서 배포한 성인지통계 관련 보도자료를 제공합니다. 10개 제목 검색			
번호	제목	기관명	배포일
2156	자궁경부암, 20~30대 여성암 진료인원 중 각각 10% 이상 차지	건강보험심사평가원	2016.05.30
2155	1~12세 자녀 둔 경기도 35~44세 여성, '시간제 일자리 희망'	경기도가족여성연구원	2016.05.27
2154	인구통합(2016년 3월)	서울통계	2016.05.26
2153	2016년 3월 경제활동인구조사 근로형태별 부가조사 결과	통계청	2016.05.26
2152	2016년 3월 인구통합(출생, 사망, 혼인, 이혼)	통계청	2016.05.26
2151	전역그래프스(2016.5)	서울시여성가족재단	2016.05.25
2150	2016년 통남권(부산·울산·경남) 청소년 통계	통남지방통계청	2016.05.24
2149	'알레르기성 결막염, 10대 미만 연령대가 가장 많아	국민건강보험공단	2016.05.23
2148	2015년 지역사회건강조사 주요결과 발표	보건복지부	2016.05.17
2147	강원도 여성가족부와 주력	강원도여성가족연구원	2016.05.16
<< < 1 2 3 4 5 6 7 8 9 10 > >>			

포털명	한국여성정책연구원 성인지통계 정보시스템	http://gsis.kwdi.re.kr
------------	--------------------------	------------------------

2016년 06월 정기 EVENT
성인지통계 3.0 Quiz 이벤트

이벤트기간 2016.06.01(수) - 2016.06.30(목)

당첨자발표 2016.07.11(월) 알림 > 이벤트 > 당첨자발표에서 확인하실 수 있습니다.

경품안내
1등 - 온라인 문화상품권 5만원권 [1명]
2등 - 온라인 문화상품권 3만원권 [5명]
3등 - 온라인 문화상품권 1만원권 [30명]

6월 퀴즈이벤트 당첨자발표

문제 1

2015년 초혼 연령은 남성 32.57세, 여성 30.00세로, 1990년 대비 평균 초혼 연령이 여성은 0.00세, 남성은 4.78세 증가했습니다. 빈 칸에 알맞은 통계수치는 무엇일까요?

○ 30.96 / 4.18 ○ 28.89 / 5.17
○ 29.96 / 5.18 ○ 29.76 / 5.18

힌트
남성 : 주계별 > 가족 > 가족형성 > 평균초혼연령(성/청장구역별)

2. 개인정보의 수집 및 보유

성인지통계는 법령의 규정과 정보주체의 동의에 의해서만 개인정보를 수집·보유합니다. 연구원이 법령의 규정에 근거하여 수집·보유하고 있는 개인정보파일은 다음과 같습니다.

개인정보 파일명	보유근거	보유목적	수집방법	개인정보 항목	보유기간
한국여성정책연구원 성인지통계 원시자료 신철자	정보주체동의	원시자료 제공	온라인수집	이름, 이메일, 휴대전화번호, 소속기관, 소속부서	연구임 내 처리결과 완료까지
한국여성정책연구원 성인지통계 메일링 신철자	정보주체동의	뉴스레터 발송	온라인수집	이메일	서비스 행차시까지
한국여성정책연구원 성인지통계 이벤트 참여자 및 응답자 회원	정보주체동의	이벤트 참여 및 당첨자 발표	온라인수집	이름, 주소(우편번호), 휴대전화번호, 이메일	연구임 내 처리결과 완료까지

■ 서비스 제공

- 메인화면: 총 열 가지의 콘텐츠(사이트맵, 검색, 인포그래픽, 주요지표 바로가기, 통계db주제별 바로가기, 알림창, 성평등지수/간행물/마이크로데이터, 통계DB업데이트/보도자료, 통계리포트, 관련사이트)를 한 번에 보여 줌.
- 통계DB: 주제별, 지역별, 출처별, 국제(주제별, 기관별), 특화(여성정책연구원에서 재분석한 통계)로 나눠 제공하며, 통계청과 연계 없이 직접 입력함. 또한 새 창(마이플랫폼 등)으로 통계표가 열리는 것이 아니라 화면에 보일 수 있도록 통계표 시스템을 구축함.
- 테마통계: 성평등지수, 주요지표, 통계리포트, 도표로 보기, 지도로 보기의 탭으로 구분하였음. 성평등지수에 대해 국가, 지역, 국제별로 정보를 제시하였으며, 주요지표는 통계와 관련된 그래프, 자료해석, 의미분석, 이용상 유의점 등을 함께 제공함. 통계리포트는 1~4일 간격으로 성인지통계를 그래프와 함께 제공함.
- 마이크로데이터: 여성정책연구원에서 제공하고 있는 마이크로데이터는 '전국다문화가족실태조사(2012)' 1건이지만, 별도의 회원가입 없이 마이크로데이터를 신청할 수 있음.
- DB업데이트 & 보도자료: 매주 통계기관(통계청 등)에서 작성하는 성인지 관련 통계표가 업데이트되고, 성인지관련통계가 배포되었을 때 제목만 클릭하면 각 기관에서 제공한 보도자료를 새 창으로 연결해 줌
- 이벤트: 월마다 정기 이벤트를 시행하며, 성인지통계 문제의 답을 해결하기 위해 이용자들은 통계DB를 찾아야 함. 이는 통계청 이벤트와 유사함.

■ 특이 사항

- 성별 통계에 대한 모든 정보가 담겨 있으며, 마이크로데이터를 재분석하여 통계를 보여 줌.
- 회원가입을 하지 않아도 됨(개인정보의 수집 및 보유 참고).

포털명
서울연구원 서울연구데이터서비스
http://data.si.re.kr

[사건으로 본 서울](#)
[통계연구시리즈](#)
[주제별 데이터](#)
[조사데이터](#)
[서울연구논문 공모전](#)

사진으로 본 서울

[<](#) [||](#) [>](#)

서울연구데이터 NOW

▶ 2016-09-28 [주제별데이터] 인구변화 최신자료 5건이 추가되었습니다.

사진으로 본 서울

서울시가 1995년부터 도시경관의 변화과정을 5년마다 주기적으로 기록하여 서울경관 정책과 역사연구의 기초자료로 활용하고 있는 "서울 도시형태와 경관" 사진을 제공 합니다.

서울의 도심지변화
9개의 구역별로 구성된 그림입니다.

서울은 A구역의 소멸과 변화에 발맞춰 변화하고 발전하는 도시이다. 서울이 주요 중심자연역사도시, 강남, 마포도, 영등포와 더불어 참성과 상영 등 다양한 대표적 자연유산지는 서울이시..

[영등포](#)
[발산역사](#)
[청량리역](#)
[서울월드](#)
[영선역](#)
[한남역 주변](#)
[영등포역](#)

동대문운동장 DDP

롯데월드공원

여의도공원

경향신문광장

통계연구시리즈

서울연구원에서는 각종 정책관련 자료를 지도와 그래프로 시각화 한 지도로 본 서울, 통계로 본 서울, 지표로 본 서울 등의 통계연구시리즈를 진행하였으며 각 연구결과에 대한 원문, 도표, 분석결과를 제공합니다.

토지이용
지도로 본 서울 > 2000 ~ 2012 토지이용

현 도시가 토지이용 현황을 파악하는 것은 도시계획을 위해 매우 중요하며, 서울에서는 관공의 고려한 도시관리기 기조로서 1999년 도..

[더보기]

도로시설
지도로 본 서울 > 2003 ~ 11월 교통

서울에 도로망은 방사선, 거점선, 순환형 도로망이 혼재된 형태이며, 이는 도시의 일상생활과 경제에 영향을 미친다. 1990년대에는 교통은 안개만..

[더보기]

중·고등학교
지도로 본 서울 > 2007 ~ 4월 중·고 학교 위치

서울시에 중·고등학교 수는 2000년대 현재까지 367개, 2005년 이후, 1970년대부터 꾸준히 증가하다가 2000년 이후에는 증가추세로..

[더보기]

주요 지하철 교통량 및 통행속도
지도로 본 서울 > 2003 ~ 2013 교통

보도교로는 도시 환경에서 가장 많은 교통량을 싣는, 서울시에서 대중교통의 중요한 역할을 담당하는 교통수단(차량) 및 안전교량의 건수는 교통안전에 큰 영향을 미친다.

[더보기]

주제별 데이터

인구구조

주택건축

산업경제

**토지이용
도시재생**

교통

안전환경

생활복지

DATA The Seoul Institute
GO - KR

Copyright 2016. All Rights Reserved.

서울연구데이터 서울 열린 데이터 광장

KOSIS 국가정보제공

대한민국 FAQ | 개인정보보호정책 |お問い合わせ | Contact Us

서울시 통합홈페이지

포털명	서울연구원 서울연구데이터서비스	http://data.si.re.kr
------------	------------------	----------------------

[사전으로 본 서울](#)
[통계연구시리즈](#)
[주제별 데이터](#)
[조사데이터](#)
[서울연구논문 공모전](#)

연구과제 설문조사
서울복지실태조사
서울서베이
가구통행실태조사

홈 > 조사데이터 > 서울서베이

서울서베이

■ 서울서베이 도시정책지표조사 자료

~ 저작권자명 : 서울특별시 정보기획관 통계데이터담당관
~ 출처(자료링크) : 서울 열린데이터 광장(<http://data.seoul.go.kr>)

[통계연구시리즈](#)
[주제별 데이터](#)
[조사데이터](#)
[서울연구논문 공모전](#)
[이용안내](#)

2016 제4회
2015 제3회
2014 제2회
2013 제1회

홈 > 서울연구논문 공모전 > 2016 제4회

2016 서울연구논문 공모전

서울특별시와 서울연구원은 공공데이터와 빅데이터에 대한 개발·활용을 장려하고 서울시장에 대한 학술연구를 활성화하고자 “(제4회) 공공데이터를 활용한 2016 서울연구논문 공모전”을 개최합니다. 여러분의 많은 참여를 바랍니다.

추진내용

공개자료 설명서

포털명	서울연구원 서울연구데이터서비스	http://data.si.re.kr
 「공공데이터를 활용한 2016 서울연구논문 공모전」 공개자료 안내서 (2016. 05. 03) 2016년도 서울연구논문 공모전 협력을 위한 서울시 빅데이터 캠퍼스 활용 가능 데이터 설명서 2016.4  서울연구원 The Seoul Institute 서울특별시 정보기획관 (서울시 빅데이터 캠퍼스)		

■ 서비스 제공

- 통계연구시리즈: 각종 정책관련 지표를 지도와 그래프로 시각화한 연구 결과의 원문을 제공하며, 특히 지도로 본 서울 연구에 대해서는 분석결과 데이터도 직접 내려받을 수 있도록 제공함.
- 주제별데이터: 서울시, 통계청, 국토교통부, 행정자치부 등 관련 기관으로부터 활용도 높은 데이터를 제공받아 주제별로 재정리하여 제공함. 현재 인구가구, 주택건축, 산업분야(일부)를 공개하고 있으며, 추후에 교통, 안전환경, 토지이용, 도시재생 등에 대한 데이터를 순차적으로 공개할 예정임.

■ 특이 사항

- 조사데이터: 서울연구원에서 진행한 서울서베이, 가구통행실태조사, 도시생태현황조사(준비 중)와 각 연구과제의 성과 자료(준비 중)를 무료로 제공함.

- 서울연구논문 공모전: 공공데이터와 빅데이터에 대한 개방·활용을 장려하고 서울시청에 대한 학술연구를 활성화하고자 2013년부터 시작하여 현재 4회를 맞이하였고, 작년에 한국보건사회연구원 한국의료패널과 환자조사의 공공데이터를 제공한 바 있음. 4회는 서울시연구원 제공 가능 데이터, 서울시 빅데이터캠퍼스 활용 가능 빅데이터, 서울시 정책사례 데이터 목록 등 더욱 다양하고 방대한 데이터를 가지고 논문 공모전을 진행함.

포털명

산업통상자원부·산업연구원

산업통계 분석시스템(ISTANS)

<http://istans.or.kr>

ISTANS

로그인

회원가입

공지사항

산업발동계

주제발동계

동상·무역동계

ISTANS 보고서

산업분석정보

ISTANS 소개

ISTANS Mobile App

우리의 스마트한 미래를 위한 최고의 산업통계 정보입니다. ISTANS Mobile로 지금 검색해보세요.

~ 전체검색

검색어 입력

검색

산업발동계

다양한 산업통계 데이터를 주요 산업별로 확인하실 수 있습니다.

40대 제조업

20대 서비스업

주제발동계

산업통계 주제별 분석 결과입력하여 기초하여 주제별로 확인하실 수 있습니다.

거시지표

국내총생산

시장규모

생산

물가

노동·고용

R&D·기술

에너지

물가

환경

기업경영

생산성·경쟁력

산업발전동향

산업분석정보

산업 주요 통계자료 모음

통계원천

통계 2,759건 (2016년 09월)

동상·무역동계

합리적 정책수립 및 성과측정을 지원하는 동상·무역분야의 핵심 정책 통계자료 모음

거시지표

생산

수출·수입

R&D·기술

물가

환경

ISTANS 보고서

ISTANS 보고서

ISTANS Insight

중소기업 동향지표

산업 동상·거주동계

산업발전지수

중소기업통계를 국가별로 비교하여 확인하실 수 있습니다.

성과

취업

주요

시장

Insightful Graph & Table

논문·경진대회

ISTANS 소식

ISTANS 보고서 '산업 동상·거주 동계' 제37호 2016년

2016.09.26

ISTANS 보고서 주요산업동향지표 제29호에 발간되었습니다.

2016.09.26

2016년도 ISTANS 논문·경진대회 수상작 발표 및 경기 안내

2016.09.26

2016 ISTANS 논문·기술·발명 및 신기술

2016.09.26

산업별통계

ISTANS에서 제공하는 다양한 산업통계 주제별 주요 산업별로 확인하실 수 있습니다.

40대 제조업

자동차

조선

항공

기타수송장비

일반목적기계

특수목적기계

방직·직조

천·섬유

금속

화학

기타제조업

가전

복합기

반도체

디스플레이

전자

40대 제조업 > 조선

산업동계

Smart KSS

인대

해당 데이터가 없습니다.

산업구조

MODEL

TECHNICAL NOTE

생산액 (십억 원, 2013)

65,066.03

산업별생산액 (2010~2014)

84.50

생산액 (십억 원, 2013)

17,209.89

중소기업생산액 (십억 원, 2013)

-

중소기업생산액 (십억 원, 2013)

121,534.00

중소기업생산액 (십억 원, 2013)

1,384.00

투자·연구개발

MODEL

TECHNICAL NOTE

설비투자액 (십억 원, 2014)

1,425.10

설비투자액 (십억 원, 2014)

162,480.00

설비투자액 (십억 원, 2014)

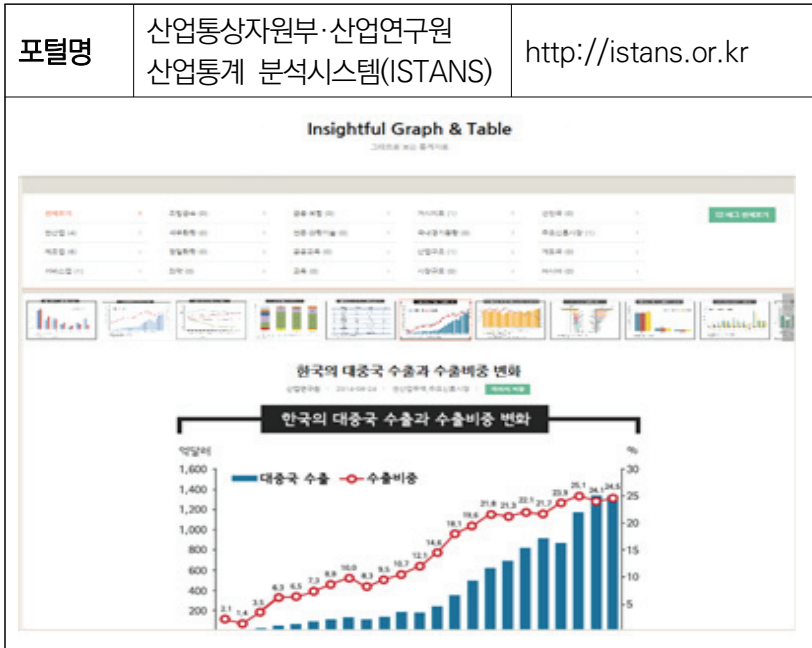
-

설비투자액 (십억 원, 2014)

-

설비투자액 (2010~2014)

설비투자액 (2010~2014)



■ 서비스 제공

- 산업별, 주제별, 미래유망산업, 통상·무역, 주요정책평가 통계로 구분하여 제공되며, 표와 그래프를 한눈에 볼 수 있음.
- 그래프로 보는 통계자료라고 하여 인포그래픽 수준은 아니지만, 지표의 그래프를 이미지화하여 제공함.

■ 특이 사항

- 국가통계포털(KOSIS)과 링크되어 있지 않고 바로 지표별 엑셀파일을 내려받을 수 있도록 구축됨.
- 논문 경진대회(공모전) 개최
- Smart RSS를 제공하여 이용자가 주로 사용하는 통계지표를 범주별로 스크랩할 수 있음.
- 통계정보 등을 확인하려면 반드시 회원가입을 해야 함.

포털명

한국보건산업진흥원
보건산업통계

<http://khiss.go.kr>

KHIDI KHISS 보건산업통계
한국보건산업진흥원 Korea Health Industry Statistics System

HOME | 사이트맵 | 보건산업주요지표 | 정부 3.0

통합검색 검색 KHISS소개

보건산업통계 산업분석정보 보고서/간행물 통계일리미 고객라운지 KHISS소개

· 의약품 · 의료기기 · 화장품 · 식품 · 의료서비스 · 보건산업총괄 · 최신산업통계 · 인가산업통계

▶▶▶ 팝업 존

대한민국 No.1 보건산업 통계포털 KHISS보건산업통계

보건산업의 흐름을 한눈에 볼 수 있는 KHISS 보건산업 통계시스템

기관별 통계 주제별 통계 산업별 통계

통계 지도 서비스
지역별 보건산업 통계정보를 이용자들이 쉽게 이해 할 수 있도록 지도 서비스를 제공 해드립니다.

보건산업 분류체계

보건산업 조사통계시스템

그레프로 보는 산업 월/분기별 통계지표

2014년 보건산업 세계시장규모(단위: US\$bn)

미국 1,027 독일 340 일본 280

· 의약품 · 의료기기 · 화장품 · 식품 · 의료서비스 · 보건산업총괄 · 최신산업통계 · 인가산업통계

뉴스식 보고서·간행물

- 의약품 허가, 전문약 높고... 2016-06-03
- 의약품 제조업 혁명자, 오... 2016-06-03
- 의료기기화 67%, "올해..." 2016-06-03
- "식면 기술기"삼균제 등 환... 2016-06-02
- 상경제약사 1분기 영업현금... 2016-06-02

보건산업 동향 **주요**
보건산업의 최신 동향과 흐름을 파악 할 수 있습니다.

보건산업 브리프
통계분석/정책제도/동향전망 정보를 제공 합니다.

글로벌 보건산업 동향
글로벌 보건산업의 최신동향과 흐름을 파악 할 수 있습니다.

KHISS 핵심지표
보건산업 핵심통계지표를 일반인부터 전문가까지 다양하게 이용 하실 수 있습니다.

한눈에 보는 지표

- 외국인 실환자수 (2014) 266,501명
- 주계인구 (2016) 50,801천명
- 원/달러 연평균 환율 (2015) 1,131.49원

KHISS Website 통계용이 사전 유권기관/관련사이트

중앙연구본부 국립암센터 한국보건산업진흥원 한국은행 한국과학기술원 보건복지부 KHIDI 한국보건산업진흥원

KHIDI KHISS 보건산업통계
한국보건산업진흥원 Korea Health Industry Statistics System

HOME | 사이트맵 | 보건산업주요지표 | 정부 3.0

통합검색 검색 KHISS소개

보건산업통계 산업분석정보 보고서/간행물 통계일리미 고객라운지 KHISS소개

· 의약품 · 의료기기 · 화장품 · 식품 · 의료서비스 · 보건산업총괄 · 최신산업통계 · 인가산업통계

보건산업통계
HEALTH INDUSTRY STATISTICS

의약품
의료기기
화장품
식품
의료서비스
보건산업총괄 >>
최신산업통계
인가산업통계

보건산업총괄 Healthcare Industry
보건산업총괄의 통계를 한눈에 볼 수 있습니다.

· 보건산업총괄

- 총괄
- 연구개발
 - **실리현황**
 - 보건의료분야 R&D 사업 - 예산현황
 - 보건의료분야 R&D 사업 - 제품개발 및 수출실적
 - 보건의료분야 R&D 사업 - 특허건수
 - 보건의료분야 R&D 사업 - 논문게제 건수
 - 보건의료분야 R&D 사업 - 연구개발 투자 현황 - 과학기술포
 - 연구개발지원
 - 특허
 - 기술수준
 - 수출입
 - 기업경영성과
 - 경제효과

· 보건산업총괄 R & D 사업 - 예산현황

- 출처정보 : 보건복지부연보 (보건복지부)
- 발표주기 : 연간
- 수록기간 : 2004~2014
- 계속여부 : 계속
- 유의사항 : 2015 보건복지통계연보(15년 11월) 기준 최신 데이터
- 연 학 처 : 한국보건산업진흥원 산업통계팀 (전화:043-713-0419)

통계요 보기

보건산업의 흐름을 한눈에 볼 수 있는 KHISS 통계서비스!

검색하기 KHISS 핵심지표 산업별통계 주제별통계 기관별통계 통계용이사전 TOP

■ 서비스 제공

- 보건산업통계: 의약품, 의료기기, 화장품, 식품, 의료서비스 등 주제별로 통계표를 제공하며 오른쪽 간편 서비스를 통해 산업별, 주제별, 기관별 통계지표를 서비스함.
- 보건산업통계 DB뿐만 아니라 산업분석정보, 보고서/간행물, 웹진 등을 발간하여 제공하며, 보건산업과 관련된 보도자료 및 통계뉴스에 대한 정보를 매일 제공함.

■ 특이 사항

- 모든 통계지표 서비스는 통계청 국가통계포털(KOSIS)과 링크로 연결함.

포털명

한국청소년정책연구원

한국 아동·청소년 데이터 아카이브

http://archive.nypi.re.kr

HOME | Sitemap | English

한국 아동·청소년 데이터 아카이브

데이터 아카이브 소개

NYPI 횡단조사

NYPI 패널조사

알림방

과학적 근거에 기초한
청소년 정책

한국청소년정책연구원이 앞장 서겠습니다.

한국아동·청소년패널조사(KCYPS)

2010년 기준 초1, 초4, 중1 재학생들을 대상으로 7개년(2010~2016)에 걸쳐 실시되는 종단조사

조사개요

조사표/데이터/코드북

학술대회 자료집

한국청소년패널조사(KYPS)

2003년 중2, 2004년 초4 재학생들을 대상으로 2008년까지 실시된 종단조사

조사개요

조사표/데이터/코드북

학술대회 자료집

NYPI 횡단조사

한국 아동 청소년의 현황과 실태를 파악할 수 있는 기초 통계자료

2014년 조사

2013년 조사

2012년 조사

2011년 조사

공지사항

주요연구결과

[안내] 한국 아동 청소년 인권실태조사, 다문화청소년 종단조사.. 2016-04-29

[안내] 한국아동-청소년패널조사 제6차년도 조사데이터 공개 일명.. 2016-04-22

[안내] 한국아동-청소년인권실태조사 2015년도 조사데이터 공개.. 2016-04-22

[알림] 한국아동청소년패널조사(KCYPS) 1-5차년도 조사데이.. 2015-11-13

[안내] 제5회 한국아동-청소년패널 학술대회 개최 2015-10-22

한국청소년연구원

WARDY
The World Association Research and Development for youth

여성가족부
Ministry of Women and Family Affairs

뷰이다운로드 | 개인정보처리방침 | 이메일무단수집거부

한국청소년정책연구원

서울특별시지치서 시청대로370 세종국책연구단지 사회정책동(ID동) 한국청소년정책연구원 6/7층

TEL : 02-551-1631 FAX : 02-551-7265-7376

청소년 관련기관 사이트

BULLETIN BOARD

알림방

공지사항

주요연구결과

데이터 활용논문

Q&A

데이터 활용논문

한국청소년정책연구원이 청소년의 밝은 미래를 열어갑니다.

구분선택

선택

검색

전체 968건, 현재 페이지 1/97

계재년도	구분	논문제목	저자명	학술지명
2016	아동청소년 패널조사	성별에 따른 아동기 사회적 관계가 자아개념 및 학교성 취향에 미치는 영향	김희영, 박수환, 양미경	한국교육문제연구, Vol.34 No. 1
2016	아동청소년 패널조사	부모의 과잉간섭이 초등학교의 휴대전화 의존에 미치는 영향: 사회적 위축과 우울의 매개효과	최태산, 김가경	청소년학연구, Vol.23 No.3
2016	아동청소년 패널조사	우리나라 아동들의 행복과 우울수준	조한익	아동교육, Vol.25 No.1
2016	아동청소년 패널조사	청소년의 오프라인 비행과 온라인 비행의 발달적특성 및 영향요인 탐색	최지영, 김재철	아동교육, Vol.25 No.1
2016	아동청소년 패널조사	초등학교의 주관적 체중과 성취도와 학교생활의 관계 : 자아존중감과 자아탄력성의 매개 효과 검증	강익영, 조전성	한국초등교육학회지, Vol.22 N o.1
2016	아동청소년 패널조사	The Factor of Longitudinal Change in Adolescents' Multicultural Acceptance	Jaeeun Roh, Jin- Kee Ha	다문화변천연구, Vol.21 No.

■ 서비스 제공

- 청소년정책연구원에서 수행하는 횡단조사와 패널조사에 대한 조사개요 및 마이크로데이터, 보고서 정보를 제공함.

■ 특이 사항

- 연구과제별로 추진되었던 아동/청소년 관련된 조사를 효율성 제고와 신뢰성 확보, 예산절감 등을 고려하여 통합조사로 실시함.
- 조사마다 제공되는 데이터 형태(SPSS, SAS, STATA)가 다르지만 데이터 제공이 무료임.
- 우리 원 한국복지패널, 한국의료패널 홈페이지와 유사함.

포털명

한국노동연구원 KLI 노동통계 Archive, 한국노동패널조사

http://www.kli.re.kr

KLI 한국노동연구원

통합검색

로그인 회원가입 현재위치 사이트맵 ENGLISH

발간자료

고용노동정보

알림광장

참여광장

연구원 소개

정부3.0

고용노동정보

고용노동정보 > KLI 노동통계 Archive

KLI 노동통계 Archive

KLI 노동통계 Archive는 한국노동연구원 발간물 노동통계의 일부 내용을 발췌하여 엑셀파일로 정리한 자료파일입니다.

한국노동연구원 발간물 노동통계 바로가기

RSS

피송 개시물 77개

페이지 1 / 8

번호	제목	첨부
77	'2016 KLI 노동통계,엑셀자료 - 주요경제지표	
76	'2016 KLI 노동통계,엑셀자료 - 인구 및 고용	
75	'2016 KLI 노동통계,엑셀자료 - 임금 및 노동생산성	
74	'2016 KLI 노동통계,엑셀자료 - 근로자생활	
73	'2016 KLI 노동통계,엑셀자료 - 노사관계	
72	'2015 KLI 비정규직 노동통계,엑셀자료 - 비정규직의 규모 추이	
71	'2015 KLI 비정규직 노동통계,엑셀자료 - 비정규직의 특성별 추이	
70	'2015 KLI 비정규직 노동통계,엑셀자료 - 비정규직의 임금, 근로시간, 근무기간	
69	'2015 KLI 비정규직 노동통계,엑셀자료 - 비정규직의 복지	
68	'2015 KLI 비정규직 노동통계,엑셀자료 - 비정규직의 취업사유	

KLIPS 한국노동패널조사

통합검색

로그인 회원가입 관리자 사이트맵 ENGLISH

한국노동패널조사

조사설계 및 결과

설문지

유저가이드

코드북

보고서

데이터다운로드

대한민국 노동연구 발전을 위한 귀중한 기초자료

한국노동패널조사 (Korean Labor & Income Panel Study)

<

||

>

1-18차년도 한국노동패널 학술대회

일지및장소

2016.10.19. / 서울대학교 호암고수회관 컨벤션센터

한국노동패널조사 데이터 설명회 및 기초분석 방법론 워크숍

일지및장소

2016.7.8. / 중소기업중앙회 2층 2대회의실

데이터다운로드

SMART KLIPS

공지사항

1-17차 한국노동패널조사 자...

2016.05.24

2016년 한국노동패널조사 실...

2016.04.28

한국노동패널조사 데이터...

2016.04.28

1-18차 한국노동패널 학술대...

2016.04.28

자료활용리스트

Bonggeun Kim, John Gibson and...

2016.01.06

Woojin Kang, 2015, An Investigat...

2016.01.06

Taehyun Ahn, 2015, Reduction of...

2016.01.06

Tae Hyun Kim, Eunah Hanb, 2015...

2016.01.06

FAQ

문의하기

한국노동패널조사 보고서

[기초분석보고서] 제17차(2014)년도 한국노동패널 기초분석보고서

[워킹페이퍼] 김기현 - 2015, 청년층 고용

[워킹페이퍼] 이윤희, 이예스다 - 2015, Parental Son Preference, Gender Role Attitud...

[기초분석보고서] 제16차(2013)년도 한국노동패널 기초분석보고서

■ 서비스 제공

- KLI 노동통계 Archive: 한국노동연구원 발간물 노동통계의 일부 내용을 발췌하여 엑셀파일로 정리하여 제공함.
- 한국노동패널조사: 우리 원 한국복지패널, 한국의료패널 홈페이지와 유사함.



KIPF
한국국제재정연구원
KOREA INSTITUTE OF PUBLIC FINANCE



연구발간자료
조세·재정·공공기관 정보
조세·재정·공공기관 협언
연구원 연례
정보공개

재정포럼
조세
재정
공공기관
경제·사회
자주 찾는 보고서

국제사태변동 및 지방재정경감의 국내 조세 재정 동향, IMF에서 연 2회 발간하는 World Economic Outlook 보고서는 OECD의 Economic Outlook 보고서의 주요 경제 및 재정변수 통계, OECD가 제공하는 조세관련 통계자료 등을 수록 및 국가별로 정리한 통계자료이며 주요 통계에 대해서는 일반 국민들이 보다 쉽게 언론 통계를 이해하기 용해하기 위해 인포그래픽 콘텐츠를 제공합니다.

조세	재정	해외	인포그래픽
----	----	----	-------

전체 ▶ 조세 ▶ 재정 ▶ 해외

일반회계

질문 : | 수록기간 : 2001~2014 | 자료출처 : 기획재정부, 각...
일반회계 세입세출 결산규모는 일반회계 세입세출 결산규모의 내용을 담고 있으며 세입규모는 1948년 조세, 전매납부금, 관계수입, 기타특별회계수입, 장려금, 산업보충자금, 대박금수입, 국유재산차익금, 외국국채의 항목으로 세입규모는 군사비, 비군사비, 경찰및공공소방, 산업및공공사업비, 기타의 항목으로 통계제국을 시작하여 2009년 세입규모는 조세(내국세, 관세, 방위세, 교통세, 교육세, 문화부동산세), 국공채 발행...

재정일반

질문 : | 수록기간 : 1970~2014 | 자료출처 : 한국은행 경제...
기능분류별 세입세출은 일반회계 세입 세출별 결산액, 일반회계 세출 수반별 결산액, 일반회계 세출 세출별 차역액(연도별) , 일반회계 세출 예산 기능별 분류의 자료를 제공한다.

08 지방세 제도개편 효과 분석

지방세 시도별 분포

질문 : | 수록기간 : 1966~2013 | 자료출처 : 행정안전부,...
지방세 시도별 분포는 지방세 시도별 분포를 나타내는 자료로서 7개 시도와 9개 도의 지방세액과 세액이 전체에서 차지하는 비율을 나타낸 자료로서 과거부터 최근까지의 각 지역별 지방세액과 그 비율을 변화율 살펴볼 수 있는 자료이다.

조세·재정·공공기관 정보

연론뉴스

- 조세제도
- 조세재정
- 세법 및 조세제도

재정제도

- 조세·재정·통계
- 조세·재정·공공기관 통합

국제회의

- 국제기구학술회담

승정각서
정치외교사건기

■ 서비스 제공

- 한국노동연구원 홈페이지와 마찬가지로, 한국조세재정연구원 홈페이지 내에, 조세와 재정관련 통계를 엑셀파일 형태로 제공함.
- 주요 지표에 대해 인포그래픽스를 제공함.

포털명	한국생명공학연구원 KRIIBB	http://portal.kribb.re.kr
------------	------------------	---------------------------


한국생명공학연구원정보포털

[즐거찾기](#) · [한국생명공학연구원](#)

[알림소식](#)
[하위사이트](#)
[서비스안내](#)
[자식라운지](#)
[종합상담실](#)
[포털소개](#)



바이오정보의 집합소
 세상의 모든 바이오 정보가
 한국생명공학연구원에서 제공됩니다.

[Quick Menu](#)
[분야별담당자](#)
[상담신청](#)
[자주하는질문](#)
[이용가이드](#)

바이오서비스	바이오정보	산학연협력	기타
------------------------	-----------------------	-----------------------	--------------------


국가영장류센터


미생물자원센터(KCTC)


바이오상용화센터


바이오평가센터



미생물자원센터(KCTC)
 경제협력개발기구(OECD)에서는 2030년 즈음에 세계경제에 큰 변화를 가져오는 '바이오경제시대'로 진입하게 되어 이에 따라 생명공학기술(BT)이 정보통신(IT) 등 다른 기술들과 융합하여 인류가 직면하고 있는 보건, 식량, 에너지, 환경 등 주요 난제를 해결하는데 중심적인 역할을 할 것이라고 전망한 바 있습니다. 특히 최근 생물다양성협약(CBD)과 이에 따른 '유전자원의 접근 및 이익공유'에 관한 나고야의정서(ABS)의 제정으로 유전자원의 이용에 대한 새로운 규제의 적용이 되고 있는 현실에 직면하여 선진각국은 생명산업의

[자세히보기](#)
[사이토라로가기](#)


실험동물자원센터


한국식물추출물은행


한국민간유전자은행


해독생물소재연구센터


국가연구안전관리본부

새소식

동향정보


 유전자치료제 임상연구 동향 보고서 2016-06-02
 면역계를 활용한 골절치료 기간 단축 연구 2016-06-02
 의료 로봇의 규제 : 위험 요소를 최소화... 2016-06-02
 특허 개항의 현황 및 시사점 - 사례 연구... 2016-06-02

보도자료

행사세미나

New devices, wearable system aim to predict, prevent as... 2016-06-02
 The organizer of body axes 2016-06-02
 New measurement technique shows link between T-cell... 2016-06-02
 New muscular dystrophy drug target identified 2016-06-02

궁금한 것을 물어보세요



- 전문가상담
- 문의사항
- 자주하는질문

이렇게 이용해보세요



- 이용가이드
- 개선회견

처음 방문하셨나요?



- 포털소개
- 하위사이트안내

[illegible]

■서비스 제공

○ 바이오정보에 대한 여러 사이트를 소개하는 포털로 20여 개의 생명공학관련기관의 정보를 수집하고 있으며, 생명공학관련 궁금증을 해결해 주기 위해 전문가 상담실을 운영함.

포털명

사회과학자료원

http://kossda.or.kr

HOME

Login

Membership

My KOSSDA

User Guide

Site Map

ENGLISH

KOSSDA 소개

자료검색

이용&기탁

연구&교육

이용자지원

소식

한국사회과학도서원

KOSSDA

한국사회과학자료원

아이디

비밀번호

로그인

이용자등록 | 아이디/비밀번호 찾기

통합 검색

양적 자료

질적 자료

문헌 자료

검색

분야·주제별 자료 찾기

정치/행정/법

경제/경영

사회/문화

사회문제/복지

심리/교육

북한/지역연구

주요 기관 자료 찾기

공공기관/정부연구소

민간기관/대학연구소

온라인통계분석

About

nesstar

주요 통계 사이트

사·군·구

다문화가족

EXCEL

PDF

NRF 토대연구 DB

전 자 처 리 DB

기관회원

관리자

이용자

New Users

양적/질적자료신청

문헌복사신청(DDS)

K-Poll

K-Poll은 KOSSDA에 소장 자료로 구축한 여론조사 데이터 링크입니다

질적 자료

텍스트, 동영상, 녹음, 이미지 등 다양한 형태의 지역공동체, 빈곤, 가족, 해외동포, 정치 및 사회운동 관련 자료를 소장하고 있습니다.

공지/소식

최신 자료

공 지 | 소 식

대선발전연구원, 연구보고서 20권 기증

[2016/07/27]

경기도가족여성연구원, 2013년 영유아보육 관련 조사자료 2건 ...

[2016/07/09]

성균관대 서베이리서치센터, 「한국종합사회조사」 2014년 자료 ...

[2016/07/05]

한국노동연구원, 「한국노동패널조사」 1~17차 자료 기탁

[2016/06/24]

아산정책연구원, 「월례 여론조사」 2013-2014년 자료 기탁

[2016/06/14]

과학기술정책연구원, 「한국기업혁신조사」 및 「혁신인력활동조...

[2016/06/10]

KOSSDA Consortium

한국중소산업정책연구원

한국행정연구원 - KIPA

산업안전보건연구원

EAI

Copyright (c) Korea Social Science Data Archive. All rights reserved.

서울시 중로구 연남동 1길 25 (우:110-094) Tel.02-767-9500 Fax.02-767-9595 E-mail

이용약관 | 기탁약관 | 데이터다운로드 | 보도자료 | 뉴스센터 | 이인요계단

nesstar

DESCRIPTION TABULATION

English

Korean

한국중소산업정책연구원(KOSSA)

1999 로동조사

경제 위기조사 2차

한국 노동조사

가계 조사

노년 조사

청소년 조사

교육 관련 조사

행정연구 개발 조사

교육 및 지방행정 관련 조사

군복자 생활 및 인식 조사

마주노동자 조사

노년층 조사

민사 및 조직관련 조사

기업 및 사업장 조사

산업안전 및 보건 관련 조사

경제위기 조사(1997년 2차)

문화체육 관련 조사

한국노동조사

한국노동패널조사

교육 관련 조사

한국노동조사

행정 관련 조사

지역공동체(지역사회) 관련 조사

노년 및 복지조사

삶의 질 조사

자주 조사-농촌 조사

기타 조사

Abstract:

「영·소년 자녀와 어머니 관계에 관한 조사」는 부모자녀 간 관계를 자립교한 관점에서 살펴보는 시도이며 어머니가 자녀로부터 제공하는 지원과 자녀 어머니로부터 제공받는 지원을 비교하기 위한 연구자료를 수집하고자 시행한 것이다.

조사는 만 24세 이하의 청소년과 중·고등학교의 자녀를 둔 어머니를 대상으로 두 차례에 걸쳐 실시되었다. 이 자료는 각 조사에서 자녀 지원을 대상으로 수집한 것으로 그 설문지는 서비스, 물질, 금전, 정보, 지력, 애정 등 6가지 차원영역에 걸친 어머니의 지원내용 실태와 요인관계 면에서 관한 문항들로 구성되어 있다.

View All Studies

Copyright © 2001-2006 NSD - All Rights Reserved - Contact Nesstar Support

KOSSDA

■ 서비스 제공

- 정치, 경제, 사회, 심리, 문화 등 사회과학 전 영역에 걸친 양적 및 질적 자료, 연구문헌을 소장하여 필요한 이용자에게 제공함.
- 연구보고서를 발간하며 이용자들에게 통계방법론 교육 서비스를 제공함(비용 지불).
- 온라인 통계분석 시스템(Networked Social Science Tools and Resources)을 제공하여 이용자들이 웹상에서 통계 분석을 할 수 있도록 구현함.

■ 특이 사항

- 다양한 양적, 질적 자료를 관리하고 제공함.
- 온라인 통계분석 시스템을 제공함.

포털명

미래창조과학부·
정보통신정책연구원 ITSTAT

http://www.itstat.go.kr



오른 ICT 통계로 활용



IT STAT IT STATISTICS OF KOREA

HOME

HELP DESK

SITEMAP

ENGLISH

가시경제

ICT산업

SW산업

통신/방송

정보화/ICT활용

R&D

우편/금융

국제ICT지수

통계자료실

통계알리미

세계 속 우리ICT

ITU ICT 발전지수 1위

UN 전자정부준비지수 1위

IMD 국가경쟁력지수 기술인프라부문 8위

WEF 네트워크준비지수 10위

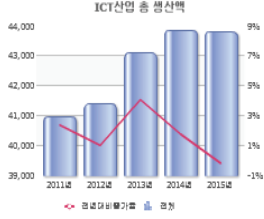
수출 상위10대 품목

01 반도체	373.67
02 휴대폰	164.81
03 평판디스플레이	164.76
04 집속부품(기구부품)	42.99
05 주변기기	39.78
06 축전기	31.09
07 TV(부품들 포함)	25.77
08 계측기	12.44
09 전구 및 조명장치	11.3
10 수동부품	10.56

(2018.1~8, 억 달러)

ICT주요통계

ICT산업 총 생산액



2011년

2012년

2013년

2014년

2015년

ICT산업 총 생산액

2011년

2012년

2013년

2014년

2015년

거시경제지표

지표	지표값
경제성장률(2016.2Q)	0.7%
1인당GDP(2015)	27,340달러
실업률(2016.2Q)	3.8%
국제유가(2016.9.26)	45.93달러
수출(2016.2Q)	1,260,091억달러
수입(2016.2Q)	989,434억달러
원/달러환율(2016.9.27)	1,106.9원

서비스 가입자 현황

이동전화	6,000만 명 (2016.7)
스마트폰	4,548만 명 (2016.7)
IPTV	1,309만 명 (2016.4)
태블릿PC	56만 명 (2016.7)
초고속인터넷	2,037만 명 (2016.7)

우리 경제 속 ICT

경제성장기여도	GDP비중(명목)	수출 비중	인력 비중
0.6%	8.9%	30.3%	9.8%
2015.4Q	2015.4Q	2014(잠정)	2013(잠정)

통계전체목록보기

Statistics Download

ICT 50대 통계지표

Statistical indicators

ICT산업 분류체계

Statistics Classification

모바일 서비스

Mobile Service

ITSTAT 이용안내

ITSTAT UserGuide

통계 알리미

2016년 8월 생산자물가지수

2016년 8월 수출입물가지수

2016년 8월 수출입 동향

최근의 국내외 경제동향(2016.9)

통계 자료실

등록된 게시글이 없습니다.

통합 검색

검색

최근검색어

2012 | 2007 | 정보격차

연구논문

정보격차 | 농어민 | 스마트

분

KOSIS 국가통계포털

ECOS 한국은행 경제통계시스템

나라지표

KISA ISIS 인터넷통계정보검색시스템

인터넷통계대응센터

KISDI

미래창조과학부

KISDI 정보통신정책연구원

ITPIND

(27672) 충청북도 진천군 덕산면 정봉로 18 TEL : 043-531-4114

Copyright © Ministry of Science, ICT and Future Planning. All Rights Reserved.



■ 서비스 제공

- ICT에 대한 주요 통계를 한눈에 파악할 수 있도록 미래부가 정보통신 정책연구원을 통해 운영함. 거시경제, ICT 산업, SW 산업, 통신·방송, 정보화·ICT활용, 연구개발, 우편금융, 국제ICT 지수 등에 대한 통계를 제공함.

■ 특이 사항

- 최근 자료 갱신이 늦다는 지적을 받음.

포털명

한국개발연구원 경제정보센터

http://eiec.kdi.re.kr

+

로그인 / JOIN

HOME

경제정책정보

경제교육

경제정보센터

Family Site

사이트를 선택하세요

개인경제교육팀 | 이원석

이원석@kdi.re.kr | 교육지원팀

SDW@

세종특별자치시 남세종로 202 (한국은행국가개발연구원)

Tel. (043)533-4114

COPYRIGHT © 2014 KDI DEVELOPMENT INSTITUTE ALL RIGHT RESERVED

최신자료

경제정보센터 최신소식 및 한 자리해

일방

주거자금대출

및 끌어가는 시스템으로 가계부채 구.

우리나라 가계부채는 지난 3월 말 기준 1,105조원에 육박했다. 그중 절반 이상이 주택을 담보로 해 자금만 주택 담보대출이다. '집이라는 단어의 두 차이를 바꾸면' 등.

공지사항

[06/27] 국가정책개발 2기(일교식 및 총회)

[06/03] 경제교육 표준교과서 교사용 지도서 발간

[06/15] 경제개발 역사포럼(2015년 5월)

[05/15] 경제개발 역사포럼(2015년 4월)

[06/30] 2015학년도 학계 경제수업 마스터

경제정책정보

경제부처 및 민간기관의 정책제출자료 및 보도자료

경제정책자료

2015년 제17차 "경제관계장관회의" 개최

"민간서민금융회사 역할 강화방안, 추진"

2015년 상반기 기업활동 동향 및 주요동향

2015년 8월중 금융시장 동향

2015년 8월 수출입물가지수

국내외 연구자료

노동시장 개혁이 주요 경제 점검

근로자 관측수령 및 임금격차 분석

2015년 상반기 국내은행 경제정보 분석결과 시사

ICT 산업의 발전 과정과 시사점

경제교육

경제교육의 핵심 - 경제교육부, 참여프로그램

경제교육

경제현장

경제교육

경제교육

경제교육

경제교육

경제교육

경제교육

경제정책정보

HOME

경제정책자료

국내연구자료

국외연구자료

수정별 목록

발간일별 목록

소개

경제정책자료

경제정보센터 및 소속기관 국내외 유수기관에서 발표된 보도자료 및 정책자료들을 매일 수집하여 신속히 제공하고 있습니다.

Hot Issue

핀테크산업 육성

핀테크 산업 육성 및 스타트업 육성 지원

국내 핀테크 기업의 해외진출 지원방안

핀테크 산업 활성화를 위한 단계별 추진전략과 향후 과제

시간선택제 도입

시간선택제 2년의 시간을 이야기합니다 (고용노동부)

277개 분야에 시간선택제 도입 추진 (고용노동부)

"시간선택제 일자리 창출효과" 2015년도 여파 분석

환율대책

G20 금융물안과 세계경제 회복을 위한 정책공조에 합의

2015년 5월 소비자물가 동향

2015년 24분기 가동률 조사

추천자료

제6회 글로벌 경제, 3대 불확실성 확대 로인 / 현대경제연구원

제2016년 국가경쟁력 진단 / 한국경제연구원

글로벌 경제에 대한 전문가의견 조사 / 대한상공회의소

제6회 60세 이상의 기업가정신 조사 / 대한상공회의소

한국경제연구원의 "우리경제 원동력 평가 및 대책" 설문조사 결과 / 한국경제연구원

동향, 전망

"OECD Economic Outlook" 발표

2016년 4월 소비자물가 동향

2016년 5월 소비자물가 동향

2016년 5월 수출입 동향

2016년 5월 소비자물가동향

법, 제도

소비자 기본법 개정안 입법예고

도론산업 활성화 등 정책적 규정

"주" 중등교육법 시행령, 입법예고

환경기술평가사업 관리 대안 강화방안

중소기업 대우로 산업단지 개발 절차

"소방차 통행가능 의무화" 아라트 전

동계

2016년 1/4분기 국민소득(잠정)

2016년 4월 소비자물가 동향

2016년 5월 소비자물가 동향

2016년 4월 국제지지(잠정)

2016년 5월 수출입 동향

2016년 5월 소비자물가동향

국민과 함께하는 규제개혁

규제정보포털

포털명	한국개발연구원 경제정보센터	http://eiec.kdi.re.kr
------------	----------------	---



15개 경제부처가 만드는
국내 유일의 경제 정책 정보지

게재 월호 상세검색

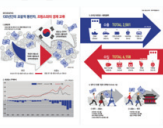
SPECIAL THEME ISSUE INTERVIEW 집중토론 Now 경제정책해설 세계는 지금 COLUMN **▶ INFOGRAPHIC** 독자 메뉴

HOME > INFOGRAPHIC

INFOGRAPHIC

프랑스와의 경제 교류 현황

2016년 06월호



금융개혁 1년, 성과와 과제

2016년 05월호



아베노믹스의 경제성과

2016년 04월호



4차 산업혁명과 일자리의 미래

2016년 03월호



저출산·고령화 대한민국, 2020년 달라진다

2016년 02월호



2016 세계경제 한국경제

2016년 01월호



지방자치 20년의 변화와 성과

2015년 12월호



20대, 하루 평균 3시간 44분 모바일 이용

2015년 11월호



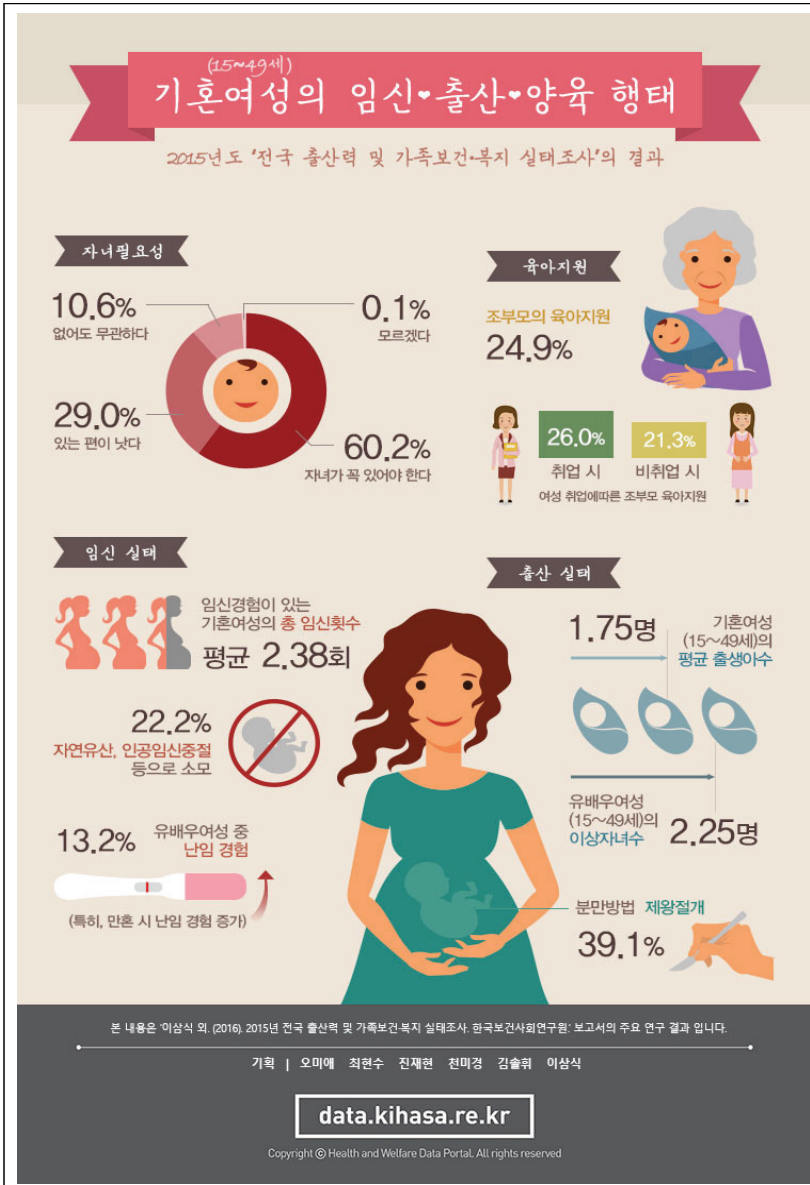
■ 서비스 제공

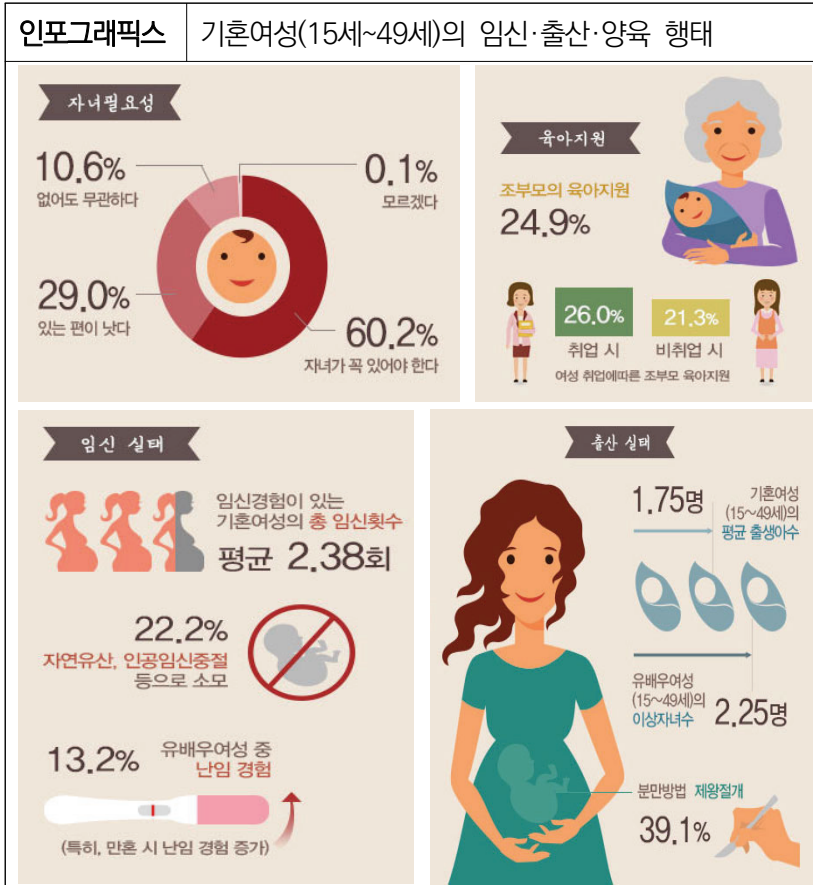
- 경제정책정보: 경제관련 부처 및 소속기관에서 발표하는 보도자료 및 정책 해설자료와 국내외 우수기관에서 발표한 연구자료들을 매일 수집·선별·분류하여 DB에 구축하고 홈페이지, 이메일서비스, 모바일 등 온라인 정보화 수단을 이용하여 실수요자에게 신속하게 제공함.
- 경제정보: 월간 『나라경제』와 Economic Bulletin 등 정기 간행물을 제공함.
- 경제교육: 『click 경제교육』을 발간하여 전국, 고교생 경제한마당, 교사 경제연수, 시민사회 경제교실, 청소년 경제교실 등 다양한 경제교육에 콘텐츠를 제공함.

■ 특이 사항

- 경제 통계 관련 자료는 경제정책정보에서 경제부처 및 관련기관의 정책해설자료 및 보도자료를 제공함.
- 한국개발연구원(KDI)은 15개 경제부처가 만드는 국내 유일의 경제 정책 정보지 '나라경제' 간행물을 월간으로 제공하고 있음. 인포그래픽으로 이용자에게 신선하게 정보를 전달함.
- Chrome 브라우저에서는 문제가 없지만, Explorer 브라우저 속성에 따라 호환성 문제가 발생함.

부록 2. 보건복지데이터포털 제공 인포그래픽스





■ 주요 내용

- 결과 배치: 소주제별(자녀필요성, 육아지원, 임신실태, 출산실태) 주요 문항들에 대한 응답 배치에서 저작물 중심에 임신부 그림을 오게 하여 기혼여성에 대한 조사결과물임을 표현함.
- 시각화: 임신부, 태아, 아이를 안은 할머니, 임신테스트기 등의 연상 이미지를 활용하여 조사결과표보다 쉽고 빠르게 정보를 전달함.

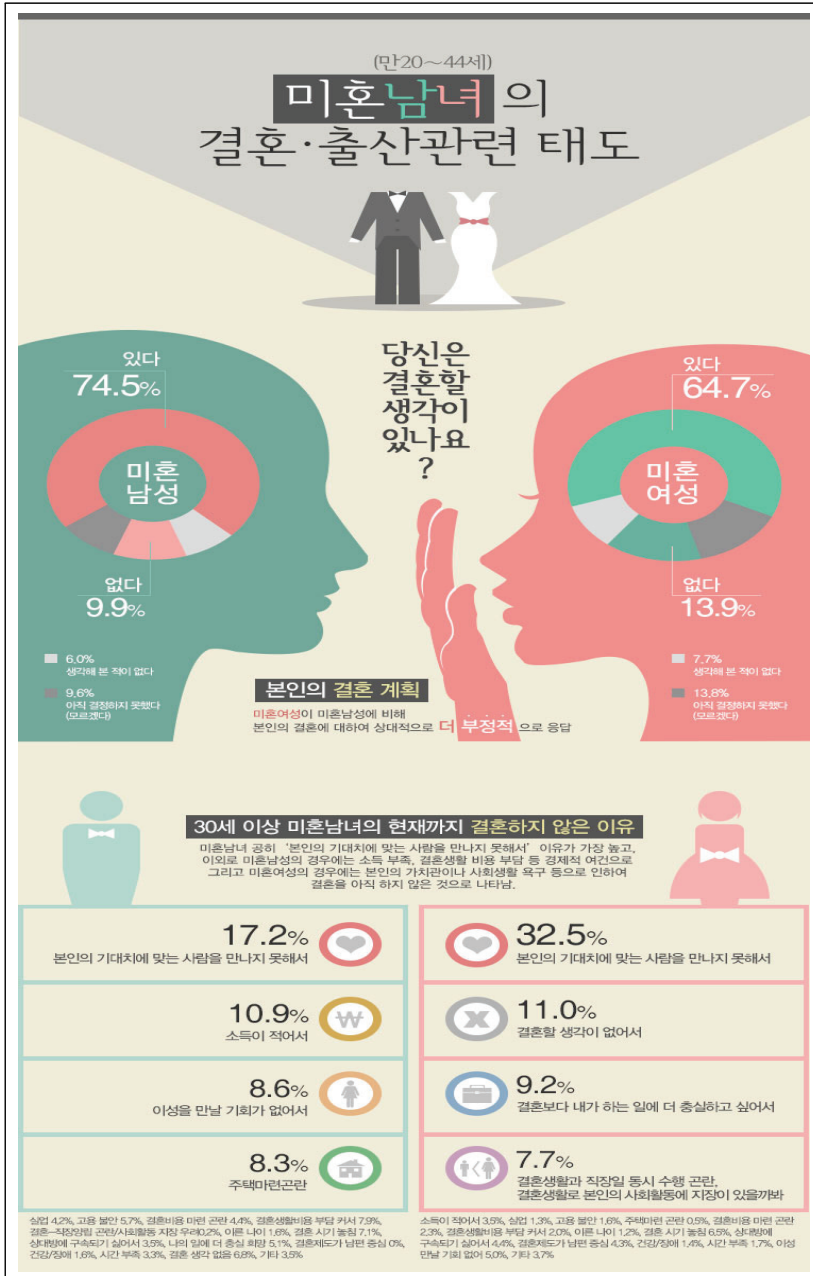


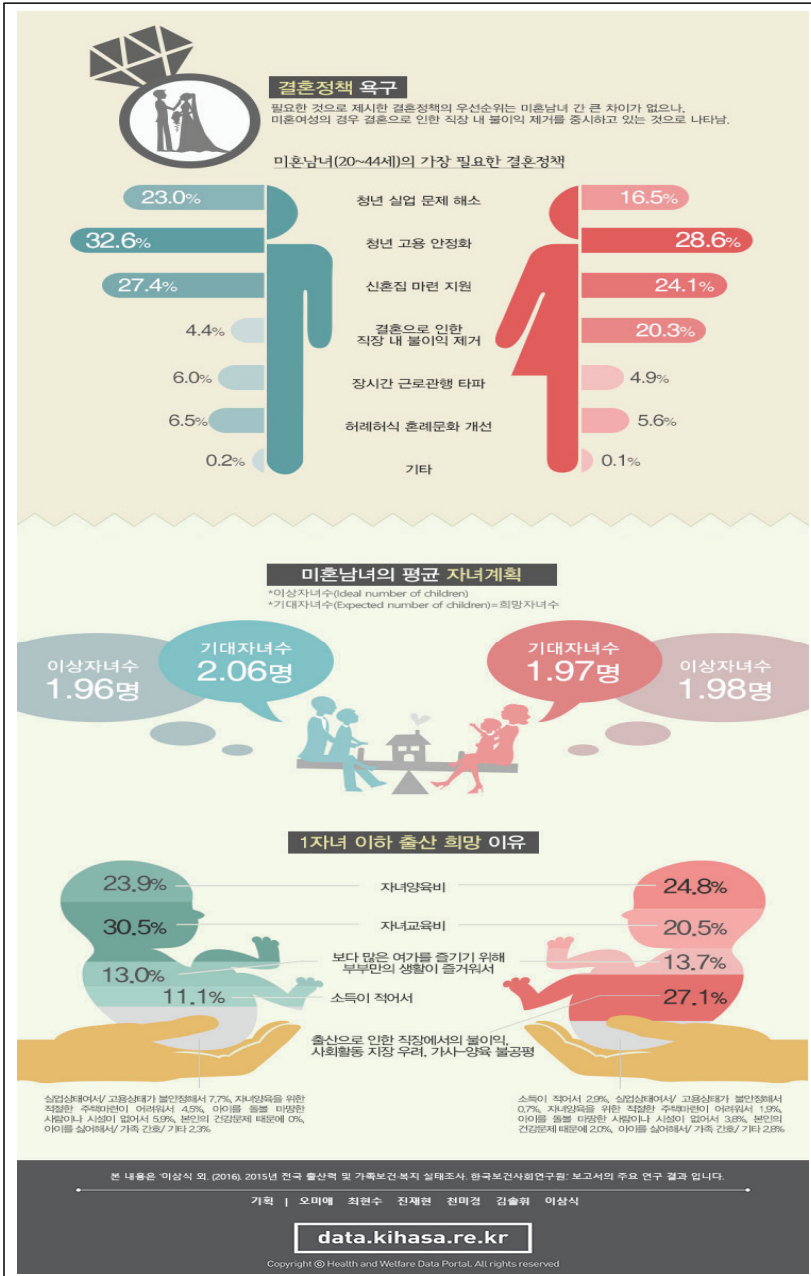
■ 자료 출처

- 활용자료 표시: 부제에 조사이름을 기재하고, 저작물 하단에 “이삼식 외. (2016). 2015년 전국 출산력 및 가족보건·복지 실태조사. 한국보건사회연구원.” 보고서의 주요 연구 결과임을 밝힘.
- 보건복지데이터포털의 주소와 해당 저작물의 기획자를 제시함. 보건복지데이터포털사이트를 홍보하고, 담당자에게 관련 내용을 문의할 수 있음.

■ 저작권

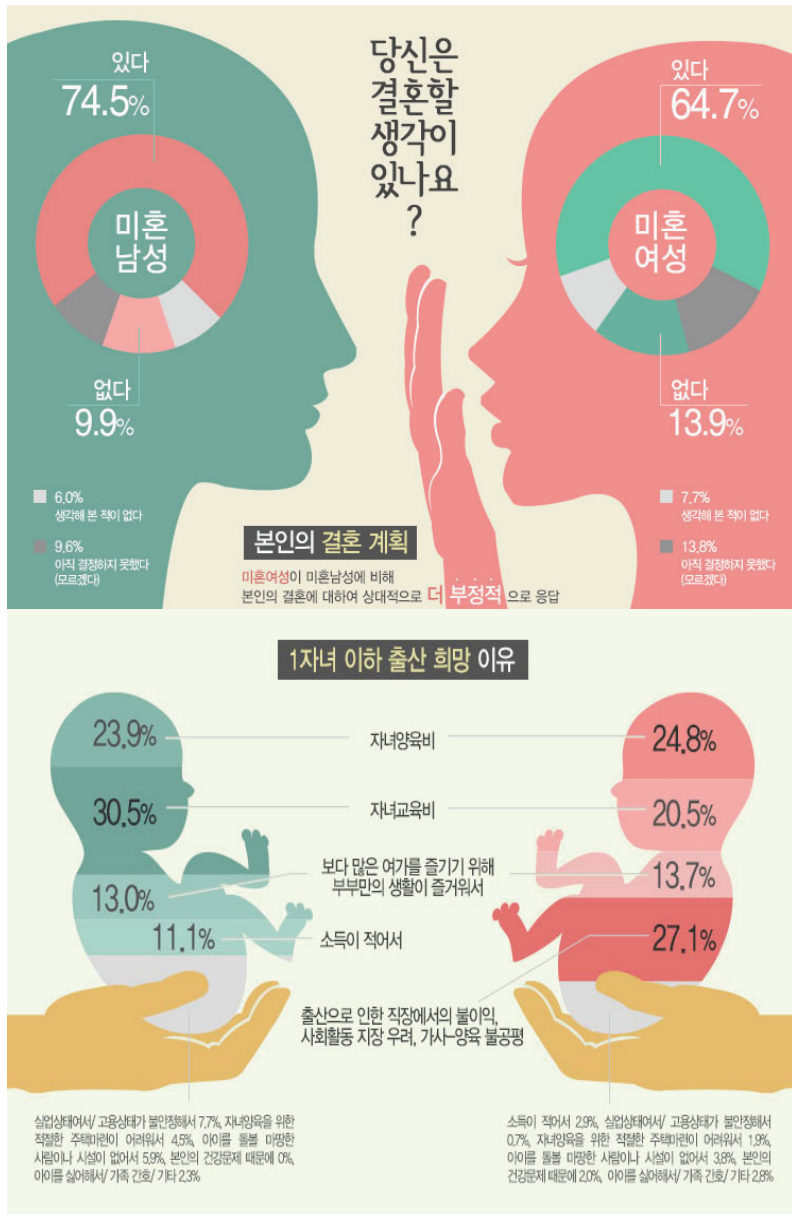
- 해당 저작물의 저작권이 보건복지데이터포털에 있음을 하단에 제시하고, 비영리적 목적에 한하여 이용자들이 재사용할 때 출처를 제공함.

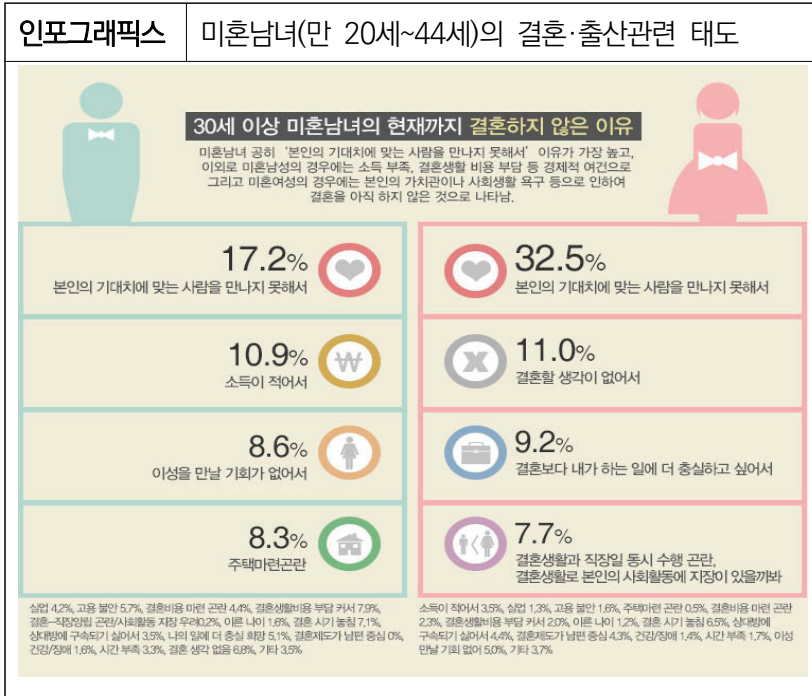




인포그래픽스

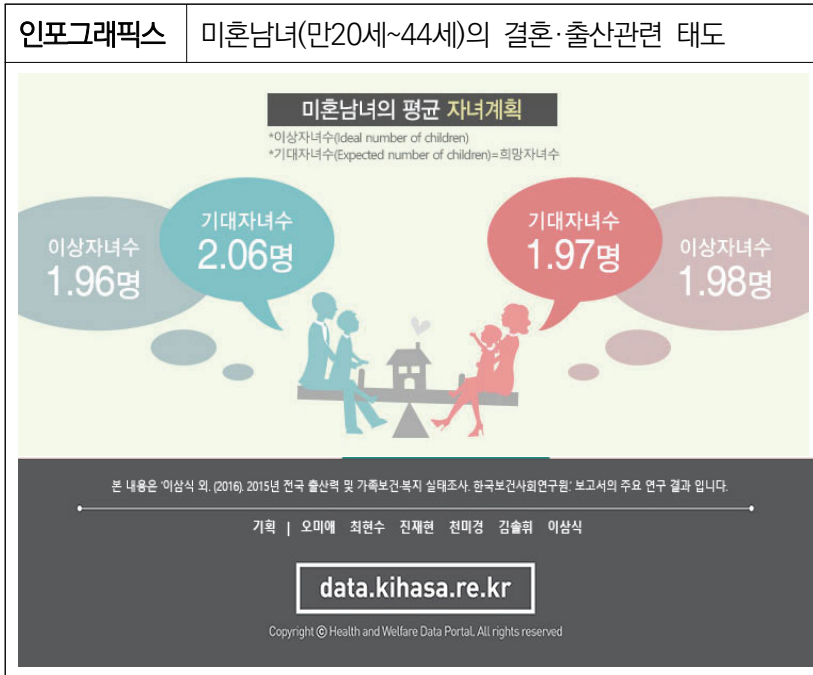
미혼남녀(만 20세~44세)의 결혼·출산관련 태도





■ 주요내용

- 전체 저작물의 원편과 청색은 남성의 응답, 오른편과 적색은 여성의 응답으로 통일성 있게 구성함.
- 조사결과의 소주제별 응답내용을 슬라이드형으로 분할하여 결혼계획 부터 1자녀 이하 출산 희망 이유까지 시간순으로 전개함.
- 부문별 요약사항을 소주제 제목 아래에 기재함(본인의 결혼계획, 현재 까지 결혼하지 않은 이유, 결혼정책 욕구 요약).
- 보기항목을 메타포(metaphor)로 표현함(예: 주택마련곤란=집 그림).
- 시각화로 표현한 주요결과 이외의 세부사항을 부문별 하단에 문자로 제공하여 정보 누락을 방지함.



■ 자료 출처

- 활용자료 표시: 저작물 하단에 “이삼식 외. (2016). 2015년 전국 출산력 및 가족보건·복지 실태조사. 한국보건사회연구원.” 보고서의 주요 연구 결과임을 밝힘
- 보건복지데이터포털의 주소와 해당 저작물의 기획자를 제시함. 보건·복지데이터포털사이트를 홍보하고, 담당자에게 관련 내용을 문의할 수 있음.

■ 저작권

- 해당 저작물의 저작권이 보건복지데이터포털에 있음을 하단에 제시하고, 비영리적 목적에 한하여 이용자들이 재사용할 때 출처를 제공함.

간행물회원제 안내

▶ 회원에 대한 특전

- 본 연구원이 발행하는 판매용 보고서는 물론 「보건복지포럼」, 「보건사회연구」도 무료로 받아보실 수 있으며 일반 서점에서 구입할 수 없는 비매용 간행물은 실비로 제공합니다.
- 가입기간 중 회비가 인상되는 경우라도 추가 부담이 없습니다.

▶ 회원종류

- 전체간행물회원 : 120,000원
- 보건분야 간행물회원 : 75,000원
- 사회분야 간행물회원 : 75,000원
- 정기간행물회원 : 35,000원

▶ 가입방법

- 홈페이지(www.kihasa.re.kr) - 발간자료 - 간행물구독안내

▶ 문의처

- (30147) 세종특별자치시 시청대로 370 세종국책연구단지 사회정책동 1F~5F
간행물 담당자 (Tel: 044-287-8157)

KIHASA 도서 판매처

- | | |
|---|---|
| ■ 한국경제서적(총판) 737-7498 | ■ 교보문고(광화문점) 1544-1900 |
| ■ 영풍문고(종로점) 399-5600 | ■ 서울문고(종로점) 2198-2307 |
| ■ Yes24 http://www.yes24.com | ■ 알라딘 http://www.aladdin.co.kr |