

**2023. 2**

**서울시립대학교**

**이용희**

<제목 차례>

[1. 연구의 배경과 목표 5](##58790610)

[2. 사회분야 통계자료 현황 6](##58790611)

[가. 사회분야 승인통계 선정 및 검색엔진 구현 6](##58790612)

[나. 통계청의 주요 승인 통계 9](##58791917)

[다. 비승인 통계자료 현황 9](##58791923)

[3. 사회분야 행정자료 현황 18](##59523362)

[가. 통계청의 행정통계DB 18](##59523380)

[나. 통계청의 민간자료DB 22](##59523460)

[다. 건강보험심사평가원DB 24](##59626542)

[라. 국민건강보험공단 코호트DB 26](##59626567)

[마. 국세청 행정자료 27](##59626571)

[바. 교육부 행정자료 29](##59627193)

[사. 사회보장정보원 DB 30](##59627323)

[아. 국가기관 보유 행정자료 DB 31](##59627331)

[4. 참고문헌 37](##784957379)

[5. 부록 38](##59627470)

[가. 승인자료 외 출처 현황 분석 38](##59627475)

[나. 통계청 행정자료 목록 69](##59627540)

[다. 건강보험심사평가원 행정자료 목록 70](##59627543)

[라. 건강보험공단 코호트 자료 78](##59627550)

[마. 국세청 DB 84](##59627556)

[바. 교육부 시스템과 DB 90](##59627562)

<표 차례>

[표 1 통계청 승인통계의 통계 분야별 개수 6](##1132533325)

[표 2 사회분야에 추가로 포함된 승인통계 7](##1132533255)

[표 3 대한민국 대표 사회 지표의 현황 11](##1132533798)

[표 4 지표 생산에 서용된 자료를 분류하는 기준 12](##1132533752)

[표 5 e-나라 지표를 생산하는데 사용된 비승인통계(일부) 13](##1133191134)

[표 6 e-나라 지표를 생산하는데 사용된 행정자료(일부) 14](##1133191136)

[표 7 e-나라 지표를 생산하는데 사용된 국외자료(일부) 15](##1133191138)

[표 8 지표 생산에 사용된 승인통계 이외의 분류별 상위 5위의 자료 16](##1133191680)

[표 9 지표 생산에 사용된 승인통계 중 상위 20위의 자료 17](##1133191684)

[표 10 통계청 사업장기초DB의 구성 항목 19](##1133368375)

[표 11 기업통계등록부 생산에 이용된 행정자료 19](##1133265211)

[표 12 통계청 민간DB 카드 소비 정보 23](##1133368378)

[표 13 건겅보험심사평가원의 보유 의료정보 25](##1133368389)

[표 14 국세청의 주요 업무 27](##1133368407)

[표 15 공공기관 행정자료 시스템 36](##1133369396)

[표 16 38](##1133369302)

[표 17 42](##1133369306)

[표 18 50](##1133369313)

[표 19 51](##1133369317)

[표 20 51](##1133369320)

[표 21 51](##1133369323)

[표 22 52](##1133369325)

[표 23 52](##1133369328)

[표 24 52](##1133369331)

[표 25 53](##1133369333)

[표 26 54](##1133369336)

[표 27 55](##1133369339)

[표 28 56](##1133369346)

[표 29 57](##1133369349)

[표 30 58](##1133369352)

[표 31 64](##1133369354)

[표 32 65](##1133369357)

[표 33 67](##1133369360)

[표 34 70](##1133369369)

[표 35 72](##1133369371)

[표 36 78](##1133369377)

[표 37 89](##1133369383)

[표 38 90](##1133369388)

[표 39 92](##1133369390)

<그림 차례>

[그림 1 1](##1859569063)

[그림 2 사회분야 승인통계 검색 도구 8](##1132533735)

[그림 3 사회분야 승인통계 편집 도구 8](##1132533737)

[그림 4 건강보험심사평가원의 행정자료 DB 관계 26](##1133368382)

[그림 5 29](##1133368415)

1. 연구의 배경과 목표
2. 사회분야 통계자료 현황
   1. 사회분야 승인통계 선정 및 검색엔진 구현

승인통계는 통계청의 승인을 받아 작성되는 공식통계로, 「통계법」 제3조 제1항 제1호에서 통계작성기관이 정부 정책의 수립·평가 또는 경제·사회현상의 연구·분석 등에 활용할 목적으로 작성되는 통계를 말한다. 통계청에서 2022년 9월 시점에 총 1293개의 승인통계가 등록되어 있다. 또한 통계청에서는 모든 통계를 30개의 통계분야로 분류하고 있다.

|  |  |
| --- | --- |
| 통계분야 | 승인통계 개수 |
| 지역통계 | 636 |
| 보건 | 63 |
| 농림 | 43 |
| 환경 | 38 |
| 노동 | 37 |
| 정보통신 | 36 |
| 범죄ﾷ안전 | 32 |
| 교통ﾷ물류 | 28 |
| 복지 | 28 |
| 기업경영 | 27 |
| 문화.여가 | 24 |
| 사회일반 | 24 |
| 국민계정 | 23 |
| 건설 | 21 |
| 도소매ﾷ서비스 | 21 |
| 정부ﾷ재정 | 21 |
| 과학ﾷ기술 | 20 |
| 광업ﾷ제조업 | 17 |
| 무역ﾷ국제수지 | 17 |
| 금융 | 16 |
| 에너지 | 16 |
| 경제일반ﾷ경기 | 15 |
| 교육ﾷ훈련 | 15 |
| 주거 | 15 |
| 수산 | 14 |
| 임금 | 11 |
| 국토이용 | 10 |
| 인구 | 10 |
| 물가 | 9 |
| 소득ﾷ소비ﾷ자산 | 6 |
| 총합 | 1,293 |

표 **오류! 일련 번호가 지정되어 있지 않습니다.** 통계청 승인통계의 통계 분야별 개수 (주황색은 사회통계 분야)

* 지역통계는 사회분야 승인통계에서 제외하였다.

승인통계 중 지역통계가 전체의 절반 정도(636개)를 차지하고 있으며 지역통계는 대부분 지방자치단체 (시도, 시군구)에서 실시되는 사회조사통계이다. 대부분의 지역통계 명칭은 지방자치단체의 기본통계, 사회조사, 경제지표조사이고 지역의 이슈와 특성을 반영하고 있다. 시도에서 작성하는 지역통계는 특정한 주제에 대한 조사통계가 있다(예를 들어 경기도아동가구주거실태조사, 경기도주요관광지방문객실태조사 등). 이러한 이유로 조사 항목이 시군구마다 다르기 때문에 통합하여 전국적 지표를 만드는데 이용하는 것이 어렵다. 본 연구에는 일단 지역통계는 사회분야 통계로 고려하지 않는다. 다만 지역에 대한 정보가 필요하면 각 지역에 대한 지역통계를 별로도 이용하는 것이 가능하다.

* 승인통계 다음과 같은 통계분야에 속하는 모든 344개의 승인통계를 우선적으로 사회통계로 분류하였다.
* 인구, 보건, 복지, 주거, 환경, 사회일반
* 과학·기술, 교육·훈련, 문화·여가, 범죄·안전,
* 노동, 도소매·서비스, 소득·소비·자산, 임금,
* 위에서 사회통계로 선정된 통계분야에 속하지 않더라도 사회영역과 관련이 있다고 판단한 다음과 같은 24개의 통계들을 사회분야로 분류하였다.

|  |  |
| --- | --- |
| 통계분야 | 사회분야 포함 승인통계 |
| 건설 | 도시정비사업현황 |
| 교통·물류 | 교통문화실태조사, 자전거이용현황, 교통접근성지표 |
| 국민계정 | 지역소득 |
| 기업경영 | 사회적기업실태조사, 일가정양립실태조사, 중증장애인생산품 생산시설 실태조사, 여성기업실태조사, 장애인기업실태조사, 소셜벤처실태조사, 소상공인실태조사 |
| 농림 | 여성농업인실태조사, 귀농귀촌실태조사, 농촌관광실태조사, 귀농어귀촌인통계 |
| 수산 | 어업인의업무상질병및손상조사 |
| 에너지 | 전력소비행태분석 |
| 정보통신 | 디지털정보격차실태조사, 웹접근성실태조사, 인터넷이용실태조사, 방송매체이용행태조사 |
| 정부·재정 | 보훈보상금지급현황, 국가보훈대상자생활실태조사 |

표 **오류! 일련 번호가 지정되어 있지 않습니다.** 사회분야에 추가로 포함된 승인통계

* 사회통계로 선정된 승인통계의 개수는 모두 368개로서 선정된 승인통계는 “통계작성의 목적”과 “조사 항목”에 대한 서술의 형식을 통일하여 통계의 명칭뿐만이 아니라 통계의 목적과 조사 항목을 이용하여 통계를 건색할 수 있는 웹기반 서비스를 구현하였다.
* 다음 그림에서 볼 수 있듯이 사회통계로 선정된 368개의 승인통계는 통계의 명칭뿐만이 아니라 통계작성의 목적과 조사 항목도 검색의 대상이 된다.

그림 **오류! 책갈피가 정의되어 있지 않습니다.** 사회분야 승인통계 검색 도구

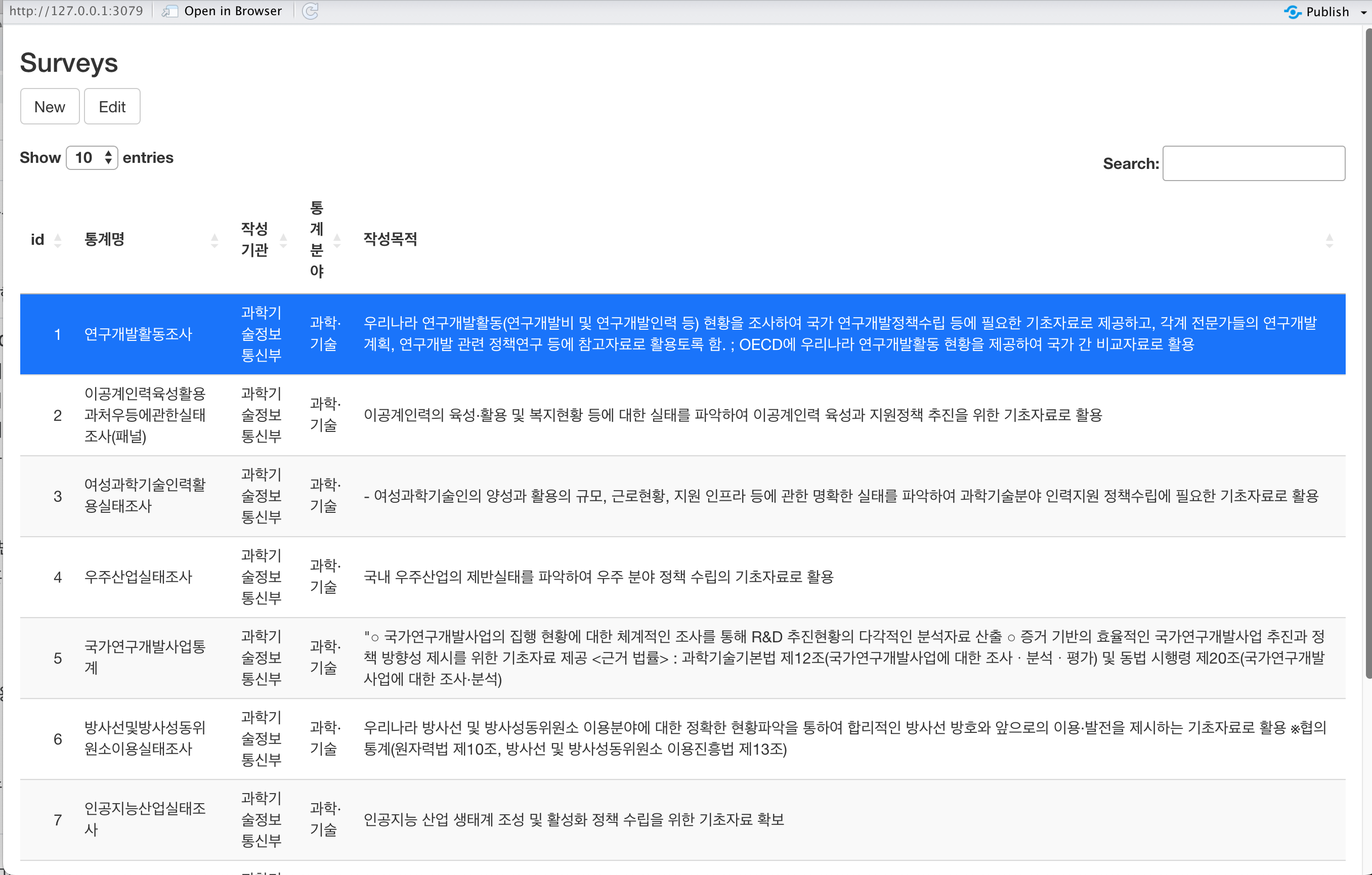
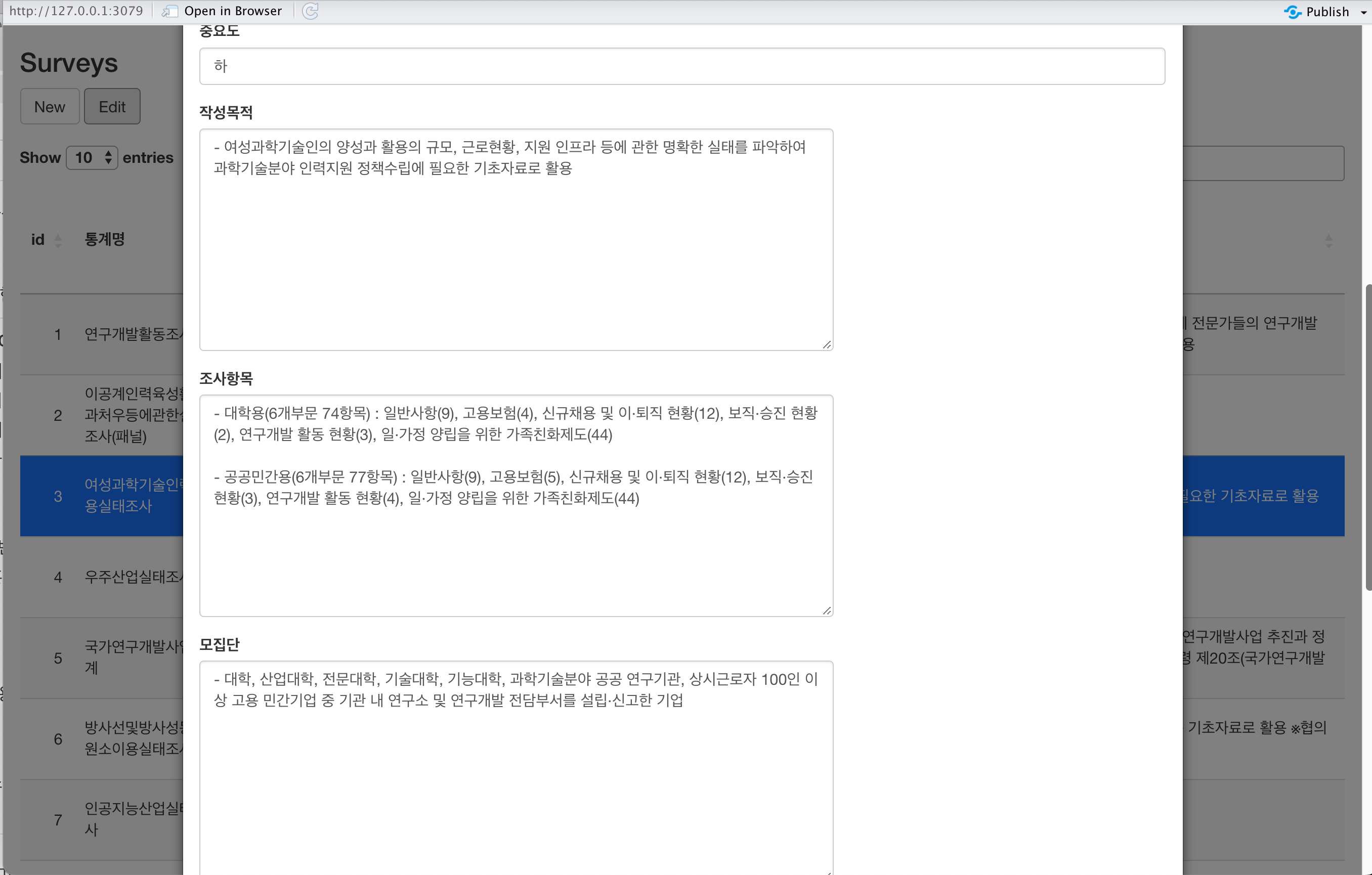
* 

그림 **오류! 책갈피가 정의되어 있지 않습니다.** 사회분야 승인통계 편집 도구

* 

* 1. 통계청의 주요 승인 통계

통계청에서 제공하는 국가통계포털(KOSIS)는 우리 나라의 주요 통계작성기관이 생상하고 있는 중요한 승인통계를 기반한 통계 서비스 시스템이다. 국가통계포털이 제공하는 통계 서비스는 통계청의 30개의 통계분야에서 지역통계, 정부재정, 물가, 환경 등을 제외한 19개 분야에 관한 통계 항목을 제공하며 통계항목의 수는 64,996개이고 사용된 승인 통계는 455개이다. 따라서 국가통계포털에 사용된 455개의 승인통계는 통계청이 발표하는 주요 지표의 기반이 되는 중요한 통계임을 알 수 있다.

* 국가통계포털에서 제공하는 통계분야는 다음과 같다.
* 인구, 사회일반, 범죄ㆍ안전, 노동, 소득ㆍ소비ㆍ자산, 보건, 복지
* 교육ㆍ훈련, 문화ㆍ여가, 주거, 국토이용, 경제일반ㆍ경기, 기업경영
* 농림, 수산, 광업ㆍ제조업, 건설, 교통ㆍ물류, 정보통신

국가통계포털에서 제공하는 다양한 통계들 중에 100개를 선정하여 KOSIS100 지표로 발표하고 있다. 아래 표는 KOSIS100 지표의 분야별 지표의 개수를 나타낸다.

* 1. 비승인 통계자료 현황

승인통계가 아닌 비승인통계의 목록을 작성하는 작업은 적절하고 유효한 데이터베이스를 구성하기 어려우므로 쉬운 작업이 아니다. 특정분야에서 승인통계가 아닌 조사통계와 행정자료를 통합적으로 관리하는 방안은 선행연구가 존재한다(오미애 와 4명, 2015; 오미애 외 5명, 2017). 사회 전 분야에서 작성되는 비승인 조사통계, 통계자료 또는 행정자료의 목록을 구하는 것은 매우 어려운 일이므로 본 연구에서는 다음과 같은 방법을 이용하여 비승인 통계자료의 현황을 파악하였다.

* 대한민국의 주요지표를 생산하는 자료의 출처를 수집하였다.

우리 나라에서 정부가 작성하는 주요지표는 다음과 같다.

|  |  |  |
| --- | --- | --- |
| **지표** | **의미** | **지표 개수** |
| e-나라지표 | - 각 중앙행정기관이 정책수립, 국정운영 점검 등에 반드시 필요하다고 인정한 각종 지표로서 국정 상황을 한눈에 쉽게 파악하고자 2006년부터 시스템을 구축하여 제공 | 531 |
| 국가발전지표 | - 국가발전지표는 국가발전의 주요 분야에 대해 합리적으로 분류체계를 작성하고, 분야별로 핵심지표를 엄선하여 구축한 지표체계  - 국가발전 상황을 점검하고 방향을 설정하기 위해 단일지표보다는 경제, 사회, 환경 등 다양한 지표로 구성된 지표체계를 통해 균형 있게 판단하려는 세계적 흐름에 부응  - 투입(input)이 아닌 성과(output) 중심의 지표  - 국가발전과 국민의 웰빙, 지속가능성을 포괄하는 지표  - 방향성(national progress)이 분명한 지표  - 정책적 노력으로 변화 가능한 지표 | 93 |
| 국민 삶의 질 지표 | - 국민 삶의 질 지표는 ’삶의 질’을 구성하는 세부 생활영역별 현황을 주요 지표로 측정 하고, 이를 통해서 국민 ’삶의 질’의 현주소를 객관적으로 제시  - 국민 삶의 질 측정은 ’삶의 질’개선을 위한 정책수립과 시행에 필요한 기초자료를 제공하는데 궁극적인 목적  - 우리사회의 장점과 취약점을 파악하고 삶의 질 증진을 위한 영역별 해결과제 및 개선방안 마련을 위한 기초자료로 활용 | 84 |
| 한국의 사회지표 | - 한국의 사회지표는 우리나라의 사회상을 종합적 집약적으로 요약하여 국민의 삶과 관련한 전반적인 경제・사회 변화를 쉽게 양적・질적으로 파악,  - 현재 사회상태를 종합적・체계적・균형적으로 알 수 있게 하는 정보의 역할 수행하며, . 사회구조 변화와 관심 분야를 파악  - 각종정책의계획수립이나정책결정및효과측정에유용한기초자료를제공 함을 궁극적인 목적으로 두고 있습니다. | 270 |
| 지속가능발전목표 | - 지속가능발전목표(SDGs)는 인류 공동의 발전을 위해 전 세계가 2030년까지 달 성하기로 합의한 정책목표(17개 목표, 168개 세부목표)로 2015년 유엔총회에서 채택됨  - SDGs 데이터 이용 활성화를 위해 지표누리로 2022년 편입함 | 207 |
| 사회보장통계 | - 사회보장통계는 사회보장 영역에 따른 분야별, 생애주기별, 정책영역(OECD SOCX)별로 분류할 수 있음  - 사회보장기본법에 따라 우리나라의 사회보장 수준 및 삶의 질과 관련된 가족, 아동, 노인, 장애인, 근로, 소득보장과 빈곤, 건강, 보육 및 교육, 문화, 주거, 에너지, 환경, 사회재정의 13개로 분류하고 있음. | 857 |
| 지표 제공 사이트와 출처:  지표누리(<https://www.index.go.kr/>), 사회보장위원회(<https://www.ssc.go.kr/main.do>) | | |

표 **오류! 일련 번호가 지정되어 있지 않습니다.** 대한민국 대표 사회 지표의 현황

위의 지표들은 각각 지표의 작성 목적이 다르며 지표 작성을 위하여 스인통계를 포함한 다양한 자료를 이용하여 작성된다. 본 연구에서는 웹크롤링(web crawling) 프로그램을 사용하여 각 지표가 발표되는 웹사이트를 검색하여 지표를 작성하는데 사용되는 자료의 출처를 수집하였다. 지표 정보의 수집 시점은 2022년 12월이다.

자료의 출처를 웹크롤링으로 검색하여 수집할 때 사용된 자료는 대부분 얻을 수 있었지만 자료를 작성하는 기관의 정보는 모든 지표에 대하여 얻을 수 없었다. 한국의 사회지표, 지속가능발전목표, 사회보장통계의 지표는 출처에 대한 정보가 일관성있게 얻어지지 않아 작성하는 기관의 정보를 파악하지 못하였다.

지표를 생산 자료의 출처에서 사용된 자료를 다음과 같이 4개로 분류하였다.

|  |  |  |
| --- | --- | --- |
| **지표 작성 자료의**  **분류 기준** | **의미** | **비고** |
| 승인통계 | 통계청의 승인통계에 속하는 통계 | 자료 출처가 명확하게 승인통계와 일치하는 경우 |
| 행정자료 | 국가 기관이 보유한 행정자료 또는  데이터베이스, 시스템에 기반한 자료 | “~현황”, “~실적”, “~신고”,  “내부자료“ 포함 |
| 비승인통계 | 승인통계 또는 행정자료로  분류할 수 없는 자료 | 통계집, 통계연감, 통계포털, 보고서 등 포함 |
| 국외자료 | 외국 국가기관이나 국제기구에서  작성 또는 제공하는 통계 자료 |  |

표 **오류! 일련 번호가 지정되어 있지 않습니다.** 지표 생산에 서용된 자료를 분류하는 기준

위와 같이 지표의 출처가 되는 자료를 분류하는 경우 다음과 같은 기준을 사용하였다.

* 출처가 정확하게 승인통계이 명칭과 일치하는 경우 자료의 출처를 승인통계로 분류하였다.
* 자료의 출처가 통계집, 통계연감, 통계포털, 통계보고서 등으로 나타나도 사실 상 승인통계로 작성된 지표일 수 있지만 지표를 생산하는 과정을 반영하기 위하여 비승인통계로 분류하였다. 또한 승인통계를 사용하면서 출처의 명칭을 정확하기 기입하지 않을 수도 있어서 출처가 비승인으로 분류될 수도 있다. 따라서 비승인통계로 분류가 되어도 자료의 실제 출처는 승인통계일 수도 있다.
* 행정자료와 비승인통계를 구분하는 기준을 정확하게 정하기 어렵기 때문에 출처에 “시스템”, “데이터베이스” 등 쉽게 행정자료임을 파악할 수 있는 경우에 출처를 행정자료로 분류하였다.
* 국외자료의 출처는 사실상 한국의 국가기관이 제공한 통계자료일 수 있지만 지표를 생산하는 과정을 반영하기 위하여 출처가 외국인 경우 모두 국외자료로 분류하였다.
* 다음 표는 e-나라지표를 만들기 위해서 사용된 비승인 자료의 일부분을 나타낸 표이다.