통계의 창

WINDOW OF STATISTICS

2025 WINTER
Vol. 36



농림어업총조사와 경제총조사

데이터로 여는 농산어촌의 내일, 2025 농림어업총조사

AI・스마트 공장 품은 경제총조사, '혁신 데이터'로 경제 대전환 시대를 열다

AI・빅데이터 시대, 5년 주기 농림어업총조사 '대변신'의 조건

경제 전체를 바라보는 통계의 창, 경제총조사의 가치와 미래

CONTENTS

통계의 창
2025 WINTER
Vol. 36

발행일 2025년 11월 30일
발행인 송성헌
발행처 통계인재개발원
기 획 차진숙, 양병렬, 현지영
주 소 대전광역시 서구 한밭대로 713(월평동) 통계센터 통계인재개발원
전 화 042-366-6131, 6134
팩 스 042-366-6498
이메일 byyang@korea.kr
디자인 및 인쇄 씨앤피(042-825-0701)

ISSN 2005-1379
ⓒ 2025. 통계인재개발원

통계 ISSUE

4 **데이터로 여는 농산어촌의 내일, 2025 농림어업총조사**
강정희 | 국가데이터처 농어업통계과 사무관

10 **AI·스마트 공장 품은 경제총조사, '혁신 데이터'로 경제 대전환 시대를 열다**
박성인 | 국가데이터처 경제총조사과 사무관

16 **AI·빅데이터 시대, 5년 주기 농림어업총조사 '대변신'의 조건**
안동환 | 서울대학교 농업생명과학대학 농경제사회학부 교수

20 **경제 전체를 바라보는 통계의 창, 경제총조사의 가치와 미래**
장우현 | 한국조세재정연구원 선임연구위원

※ '통계의창'에 실린 내용은 필자 개인의 의견이므로 필자의 소속기관이나 본지의 공식적인 견해를 대변하는 것은 아닙니다.
※ '통계의창'에 실린 내용에 대해 집필진, 운영진 외 타인의 무단 사용 및 배포를 금합니다.

당신의 땅이 농산어촌에 좋은 땅이 됩니다

2025 Census 농림어업총조사



통계 이야기

24 **데이터로 배우는 시대, 실용 통계와 소프트웨어 교육의 만남**
김갑수 | 서울교육대학교 정보교육과 교수

30 **산업 구조의 재편, 제조업이 그려가는 새로운 5년**
전상범 사무관, 김근일 주무관 | 국가데이터처 산업통계과

40 **생성형 AI 시대, 데이터 윤리의 새로운 기준**
김윤명 | 디지털정책연구소 소장

44 **인공지능 시대, 엔터테인먼트는 어떻게 진화하는가**
황서이 | 중부대학교 미디어커뮤니케이션학과 교수

48 **통계로 바라보는 세상 이야기**
신동헌 | 도서출판 지일북스 대표

52 **AI를 활용하여 텍스트 데이터 수집 및 분석하기**
구자룡 | 밸류바인 대표

데이터로 여는 농산어촌의 내일, 2025 농림어업총조사

강정희 | 국가데이터처 농어업통계과 사무관





1960년도 국세조사 조사표



1961년 농업국세조사 현장사진

농림어업총조사 연혁

센서스(CENSUS)는 완전성, 개별성, 동시성, 주기성을 준수하여 조사대상을 전수조사 함으로써 국가의 전반적인 현황을 파악한다. 농림어업총조사는 국제연합(UN) 식량농업기구(FAO)의 세계농업총조사 권고안에 따라 국제적으로 130여 개 국가가 수행하는 범세계적인 센서스이다.

우리나라는 1960년에 최초로 농림수산부에서 제1차 국세조사로 농업총조사를 시작하였고, 제1차 어업총조사는 수산청에서 1970년에 최초로 실시하였으며, 1998년에는 산림청에서 제1차 임업총조사를 실시하였다.

10년 주기였던 농업총조사와 어업총조사가 1995년부터 5년 주기로 변경되었으며, 2000년부터는 농업·어업총조사가 국가데이터처(당시 통계청)으로 이관되어 실시되었고, 2005년에는 통계청 에서 농업·임업·어업총조사를 동시 실시하였다. 1980년에 처음 시작되었던 지역조사도 2000년부터 통계청에서 실시하는 농업총조사에 포함되어 실시되었다.

2015년부터는 농림어업총조사에 인터넷조사가 도입되었고, 2020년에는 종이조사표 대신 태블릿PC를 활용한 면접조사가 시작되었다.



2005년 농림어업총조사 지침서



2010년 농림어업총조사 포스터

2025 농림어업총조사 개요

2025 농림어업총조사 실시개요

- ㅇ (법적근거) 통계법 제5조의4(총조사 실시), 통계법 제17조 제1항(지정통계), 농림어업총조사 규칙
- ㅇ (조사기준시점) '25. 12. 1. 0시 현재
- ㅇ (조사기간) 방문면접조사('25.12.1.~22.), 인터넷조사('25.11.20.~12.10.)
- ㅇ (조사대상) 조사기준시점 현재 모든 농림어가와 행정리
- ㅇ (조사항목) 133개(농림가 58, 해수면어가 30, 내수면어가 30, 지역 15)
- ㅇ (조사방법) 대면조사(태블릿PC), 비대면조사(인터넷, 종이)
- ㅇ (조사체계) 국가데이터처(주관기관), 지방자치단체(실시기관)
- ㅇ (조사인력) 18,000명 내외

2025년 농림어업총조사는 통계법 제5조의4(총조사의 실시), 통계법 제17조 제1항(지정통계 제101041호), 농림어업총조사 규칙을 근거로 하여 우리나라의 모든 농·림·어가 및 가구원의 규모, 구조, 분포, 경영형태 및 특성 등을 파악하고 통계를 생산하는 것을 목적으로 한다.

이를 위해 농림어업총조사는 국가데이터처가 주관하고 각 지방자치단체에서 조사원과 총관리자, 조사관리자, 조사지원담당자를 채용하여 현장조사를 실시하며, 조사 기간은 12월 1일 0시를 기준시점으로 인터넷조사는 11월 20일부터, 방문면접조사는 12월 1일부터 시작하여 12월 22일까지 실시한다.



2025년 농림어업총조사 추진 체계

농림어업총조사의 조사대상인 농림어가는 한국표준산업분류의 농업, 임업, 어업 중 작물재배업, 축산업, 임업, 해수면·내수면 어업, 양식업 등을 직접 경영하는 가구를 말한다. 따라서 가구가 아닌 법인·정부기관·학교·단체 등과 농림어업을 경영하더라도 농림어가 정의[1]에 부합하지 않는 규모 미만 경영 가구이거나 농림어업 관련 서비스업을 경영하는 가구는 제외한다.

또한 지역조사는 시군 조례에 따라 읍면지역에 설치된 행정리를 대상으로 조사하며 행정리

장이 조사표와 경계지도를 작성한다.

농림어업총조사의 조사방법은 인터넷조사(비대면조사)와 방문면접조사(대면조사)의 병행으로 이루어진다. 먼저 인터넷조사를 시행하고, 인터넷조사가 미완료된 가구를 대상으로 조사원이 태블릿PC를 활용하여 방문면접조사를 하게 된다.

인터넷조사는 응답자가 농림어업총조사 홈페이지[2]에서 우편 발송된 조사안내문[3]의 참여번호를 참고하여 직접 조사에 참여할 수 있으며, 조사표 입력 중 조사항목에 대한 궁금사항은 24시간 운영하는 AI 보이스봇 기반 콜센터 또는 인터넷조사 화면의 "챗봇상담"을 통해 빠르고 편리하게 해결할 수 있다.

방문면접조사는 조사원이 농림어가를 방문하여 종이조사표가 아닌 태블릿PC에 탑재된 농림어가 조사표 앱을 통해 조사표를 입력하는 방식으로 현장에서 실시간으로 항목별·항목간 오류 점검이 가능하여 오류 최소화 및 처리 시간 단축 효과를 기대할 수 있다.

농림어업총조사 준비

이번 2025 농림어업총조사의 성공적인 추진을 위해, 전주기 대비 농림어업 부문 행정자료의 활용 확대(18종->22종)를 통한 명부 작성으로 농림어가 포착을 증대하고자 한다. 또한, 농림어업의 환경 변화와 정책 수요를 반영한 조사항목을 선정하고 조사시스템 및 인프라 강화와 맞춤형 홍보를 통한 현장조사의 효율화를 통해, 농산어촌의 경쟁력 강화와 농산어민의 삶의 질 향상을 위한 정책 수립의 기초통계를 서비스하고 농림어업 부문별 실태조사를 위한 모집단을 제공하고자 한다.

이를 위해 국가데이터처는 본조사 3년전(2022년)부터 센서스 통합관리시스템, 조사모집단 및 공간정보의 구축·운영 등 조사 인프라를 체계화하였고, 조사 방식별 자료 입력 기능 및 2025 인구주택총조사의 자원 연계 활용도 철저하게 점검하고 있다.

또한 2022년부터 2024년까지 농림어업총조사 시험조사와 시범예행조사를 통해 조사항목의 타당성을 검증하고, 처 내·외 전문가와 농림어업총조사 자문위원회, 국가통계위원회[4] 상정 등 논의를 거쳐 조사항목을 확정하였다.

금주기 농림어업총조사의 조사항목은 지속가능한 농림어업 정책 수립을 위해 농산어촌의 기후변화(논벼 물관리), 영농기술 발달(스마트 농업), 인구구조 변화(국적, 고용), 행정리

1) 농가 : 논이나 밭을 1,000㎡(10a) 이상 경작 가구 또는 사육 가축 평가액이 120만원 이상 가구 또는 지난 1년간 농축산물 판매금액이 120만원 이상 가구
임가 : 산림 30,000㎡ 이상 보유하면서 지난 5년간 육림 실적이 있는 가구 또는 지난 1년간 벌목업, 양묘업을 경영한 가구 또는 지난 1년간 채취업을 포함하여 직접 생산한 임산물 판매금액이 120만원 이상인 가구,
어가 : 1년간 1개월 이상 판매목적 어로어업, 양식업 경영가구 또는 양식수산물 평가액이 120만원 이상 가구 또는 1년간 직접 어획, 양식한 수산물 판매액이 120만원 이상인 가구

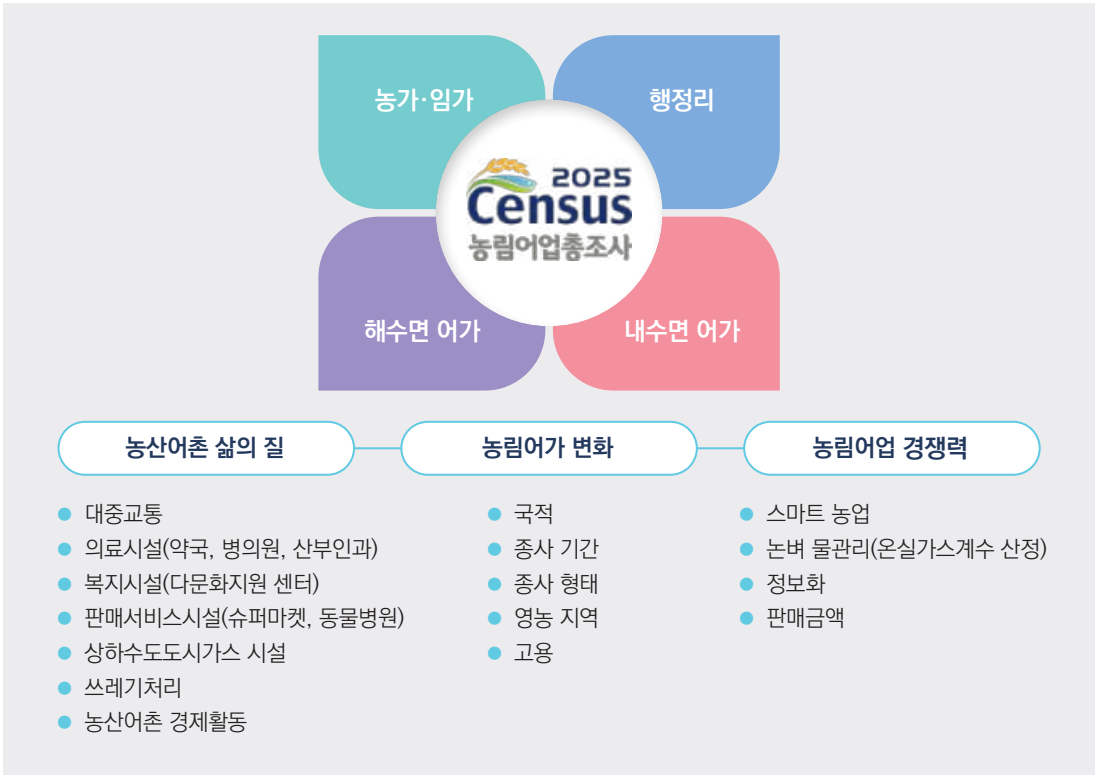
2) affcensus.go.kr

3) 농림어업총조사에 대한 안내, 참여방법, 가구별 참여번호 등이 기재되어 있으며, 우편물 수신자는 "농림어업총조사 대상 가구 귀하"로 명시할 예정임

4) 통계법 제5조의2에 의해 설치된 위원회(위원장: 부총리)로 통계의 작성·보급 및 이용에 관한 사항 심의·의결

내 생활편의 시설의 변화(산부인과, 다문화가족센터, 동물병원) 등 최근 농림어업과 농림어가의 구조 및 환경 변화를 반영하여 4종의 조사표*를 완성하였다.

* 농림가, 해수면 어가, 내수면 어가, 지역



2025년 농림어업총조사 조사항목 특징

또한 농산어촌에 거주하는 외국인의 응답 편의성을 위해 외국어 조사표를 전주기 1종(영어)에서 8종(영어, 베트남어, 필리핀어, 중국어, 일본어, 네팔어, 태국어, 캄보디아어)으로 확대하였고, 다양한 채널을 통해 대국민 및 지역 특성을 고려하여 농림어업총조사에 대한 인지도 제고(9월~), 조사요원 모집 및 총조사 실시(10월~)에 대한 대대적인 홍보를 하고 있다.





2025년 농림어업총조사 조사요원모집 포스터



2025년 농림어업총조사 실시 포스터

농림어업총조사 활용

이렇게 수집·검토·분석·집계된 농림어업총조사의 결과는 중앙행정기관과 지방자치단체 등에서 우리나라 농산어촌의 지속 가능한 미래를 설계하고 농림어업의 경쟁력을 강화하는 정책을 수립하기 위한 기초자료로 활용되고 있다.

또한 농림어업부문의 맞춤형 세분화 정책을 위한 개별 실태조사[5]의 모집단 역할을 수행하고 있다.

마지막으로, 총조사 자료는 연구자료와 학술자료, 국가 간 자료교환 및 분석자료로도 활용되고 있다.

2025 농림어업총조사 참여 요청

국가데이터처에서는 농림어업의 미래 설계와 농산어촌의 발전을 위해 농림어업총조사 결과가 농산어민에게 필요한 바른 정책으로 되돌아갈 수 있도록 온 힘을 다할 것이다.

또한 농림어업총조사에 참여한 전국 모든 농림어가의 소중한 답변은 통계법 제33조에 의해 비밀이 엄격히 보호되고, 모든 조사 내용은 통계 작성 목적으로만 사용되므로 국민 여러분께서는 안심하시고 조사에 응해 주시기를 부탁드린다.

다시 한번 2025년 11월 20일부터 12월 22일까지 시행되는 2025 농림어업총조사에 국민의 많은 관심과 적극적인 참여를 요청드린다.

5) 농·어가경제조사, 농축산물생산비조사, 농업기계이용실태조사, 여성농업인실태조사, 농업인의병무상질병손상조사, 임산물생산비조사, 임산물소득조사, 임업경영실태조사, 에너지총조사 등

AI·스마트 공장 품은 경제총조사, '혁신 데이터'로 경제 대전환 시대를 열다

박성인 | 국가데이터처 경제총조사과 사무관





경제총조사는?

경제총조사는 우리나라 모든 사업체를 대상으로 고용·생산·투입 등 산업 구조를 파악하기 위해 5년마다 실시하는 대규모 조사다. 국가데이터처와 지방자치단체가 공동으로 추진하며, 국가 경제정책과 산업구조 파악의 핵심 기반을 제공하는 경제 분야의 대표적 통계조사다.

경제총조사를 통해 동일 시점·동일 기준으로, 공식·비공식 부문을 포괄한 소지역·소규모 단위의 사업체 분포와 세부 특성 정보를 제공한다. 또한 국가 산업 구조를 이해하고 고품질 경제통계 기반을 확보하며, 지역 간 균형발전 정책 수립과 경제 변화 추이 파악·미래 예측을 지원한다.

경제총조사는 1955년부터 2004년까지 13회에 걸쳐 실시하던 산업총조사와 1968년부터 2006년까지 10회에 걸쳐 실시하던 서비스업총조사를 통합해 2011년 5월 23일 처음으로 실시하였다. 2016년과 2021년에 각각 2회 및 3회 경제총조사를 실시했으며, 2026년에는 4회 경제총조사를 실시할 예정이다. 최근에 실시한 2020년 기준 경제총조사는 행정자료 기반의 경제총조사로 모집단 범위를 확대하여 통계의 정확성을 높이고자 하였으며, 그럼에도 현장 조사를 축소하여 사업체 응답 부담을 감소시키고자 노력하였다.

경제총조사의 조사대상은 조사기준일 현재, 우리나라에서 산업활동을 수행하고 있는 모든 사업체로 단, 한국표준산업분류 21개 산업(A~U) 중 자가소비 생산 활동(T), 국제 및 외국기관(U)은 제외한다.







경제총조사의 활용

경제총조사는 대·중·소기업 간 동반성장 지원, 중소기업 경쟁력 강화, 서비스산업 선진화, 여성기업 육성·지원, 지식재산 집약 산업 육성 등 정부 정책 수립과 평가를 지원하는 기초자료로 활용되며 국민소득, 지역소득 및 산업연관표 작성, 시장집중도 분석, 사업다각화지수 등 가공통계 작성을 위한 기초자료로 활용된다.

또한 사업체노동력조사, 환경산업통계조사 등의 모집단, 광공업생산 등의 기준 연도 품목 및 가중치 산정 등 각종 표본조사의 모집단 또는 기준점을 제공하며 대학·연구기관의 연구 및 민간 기업체의 경영 실태 파악에 기초자료로 제공하여 새로운 부가가치 창출 및 상품 기획 등에 활용된다.

〈 경제총조사 KOSIS 및 마이크로데이터 이용 현황 〉

(단위: 건)

	2020년	2021년	2022년	2023년	2024년
KOSIS	167,408	270,306	296,786	277,333	344,533
마이크로데이터	1,074	896	1,411	2,326	4,037

2025년 기준 경제총조사 준비

국가데이터처는 2026년에 실시하는 경제총조사 성공을 위해 다음과 같이 체계적으로 준비하고 있다.

2023년부터 행정 자료 활용 확대, 신규 조사항목 발굴 및 조사표 개선, 온라인조사 확대 시행 등을 위한 개선방안 검토를 지속하였다. 또한, 조사항목 선정을 위해 160여 개 관련 기관 방문 등 의견을 수렴하고, 이용자 협의체·전문가 회의 및 국가데이터처·지자체 담당자 토론회를 개최하였다.

2023년 경제총조사 통합시스템 구축 정보전략계획(ISP)에 따라 2024년부터 3개년 동안 통합시스템을 구축하고 있다. 통합시스템은 AI 분류시스템 연계, 표준 에디팅 및 임퓨테이션 기능 개선 등을 통해 총조사 전 과정을 시스템화하여 효율적인 총조사 실시를 지원하고자 한다.

그리고 두 차례의 시험조사와 한 차례 시범예행조사를 통해 조사 적합성 및 타당성을 검증하였다. 2023년은 2개 시도 내 4개 읍면동 약 12,000개 사업체를 대상으로, 2024년은 2개 시군구 내 약 25,000개 사업체를 대상으로 시험조사를 실시하였고, 2025년에는 3개 시군구 내 약 40,000개 사업체를 대상으로 시범예행조사를 실시하였다. 시험조사와 시범예행조사를 통해 신규 및 개선 조사항목의 현장 적용, 온라인조사 테스트, 통합시스템 테스트 등 2026년에 시행하는 본조사 실시를 위한 각종 현장 테스트가 가능하였다.

마지막으로 2025년 기준 경제총조사의 추진 목표 및 추진 전략, 전략별 중점 추진 사항을 기본계획에 담아 국가통계위원회[1]본회의를 거쳐 최종으로 확정할 계획이다.

2025년 기준 경제총조사 추진 방향

2025년 기준 경제총조사는 '국민에게 필요한 경제총조사의 성공적 수행'을 목표로 하고, 추진 방향으로 '경제 대전환 시대의 반영 및 조사방법의 혁신'을 설정하였다. 이를 위해 다음과 같이 4개의 추진 전략과 각각의 중점 추진사항을 수립하였다.

우선 국민 수요 맞춤형 통계생산을 위한 산업생태계 변화를 반영하고자 한다. 인공지능(AI) 활용, 스마트 농장·공장, 로봇 활용, 무인매장, 외국인 종사자 수 등 산업환경 변화 및 정책 수요를 반영한 신규 조사 항목을 발굴하고, 미래산업[2] 등을 반영한 '24년 개정 한국표준산업분류(KSIC) 기준 조사를 실시하며, 전 주기 경제총조사 결과 자료에 소급 적용하여 추가 제공할 계획이다.

두 번째로 예산 절감 및 응답·조사 편의성 향상을 위한 조사 방법 효율화를 추진한다. 다양한 맞춤형 홍보와 협조공문 발송 확대 등을 통해 온라인 조사(PC, 모바일), CATI(컴퓨터를 이용한 전화조사 Computer Assisted Telephone Interviewing) 등 참여를 확대하고, 응답자 편의를 높인 온라인 조사 시스템을 구축하고자 한다. 행정자료로 대체된 유·무형 자산 항목 전면 삭제와 글자크기 확대 등의 조사표를 개선하고, 조사원 편의를 고려하여 조사표 종수를 10종에서 7종으로 축소할 계획이다. 또한, 조사원 대상 안전교육 실시 및 안전용품 지급[3] 등 안전대책을 강화할 예정이다.

세 번째로 행정자료 활용 확대를 통해 국민 응답 부담을 경감하고자 한다. 기업체 유·무형자산 항목의 국세청 법인세 자료 활용, 외국인 종사자 수 항목의 고용노동부 고용허가제 자료 활용 등 행정자료 항목 대체를 확대하고, 한국부동산원의 공동주택관리정보시스템, 한국전력공사의 전력거래금액 자료 등 행정자료를 추가 발굴하여 입수를 확대할 계획이다. 현장 규모 축소를 위해 경제적 중요성이 큰 사업체는 전수조사를 실시하고, 나머지 사업체는 표본조사(약 20%)를 실시하고자 한다. 또한 행정자료 대체 사업체[4] 등을 고려하여 현장조사 대상 사업체를 결정할 계획이다.

마지막으로 통계생산 기반 고도화를 통해 고품질 통계를 제공하고자 한다. 클라우드 기반 등 최신 IT 기술을 적용한 통합시스템을 구축하고, 표준 E&I 시스템[5] 기능 확대 등 자료처리 자동화율을 높이고자 한다. 보이스봇(챗봇) 등 AI 기반 콜센터를 도입하여 응답 정확성을 높이고, AI 기반 산업분류 자동화 시스템 활용으로 산업분류 검증 오류를 감소시킬 계획이다.



1) 통계법 제5조의2에 의해 설치된 위원회(위원장: 부총리)로 통계의 작성 · 보급 및 이용에 관한 사항 심의 · 의결2025년 기준 경제총조사 기본계획은 국가통계위원회 경제통계분과위 상정(2025. 6.)

2) 전기 자동차, 풍력발전, 이차전지 등 미래·성장산업 분류 신설 및 세분

3) 녹음기가 포함된 신분증 케이스 지급, 투넘버 서비스 가입 지원 등

4) 개인택시, 건설장비 운영업, 1인 가구 내 사업체, 무상근 사업체

5) 표준 E&I(Standard Editing and Imputation) 시스템: 조사결과의 오류나 이상치 탐색, 결측치 대체 등을 표준화된 절차에 따라 수행할 수 있도록 구축하는 시스템

2025년 기준 경제총조사 실시

[조사 개요]

□ (조사대상) 조사 기준일 현재('25. 12. 31.), 우리나라에서 산업활동을 수행하고 있는 모든 사업체(19개 산업*)
* 21개 산업(A~U) 중 자가소비 생산 활동(T), 국제 및 외국기관(U) 제외

□ (조사기간) '26. 6. 12. ~ 7. 22. (온라인 6. 1. ~ 6. 30.)

□ (조사표·항목) 7종업종별 6종, 본사 1종, 37개공통항목 11개, 특성항목 26개

□ (조사방법) 면접조사 원칙, 스마트조사온라인(PC, 모바일), 전화, 이메일, 팩스 병행

□ (조사체계) 국가데이터처(주관기관), 지방자치단체(실시기관)

□ (결과 활용) 경제관련 정부정책 수립·평가, 국민소득 등 주요 경제통계 작성, 표본조사의 모집단, 기업경영 및 연구 자료 등

이번 경제총조사의 조사대상은 2025년 12월 31일 현재 우리나라에서 산업활동을 수행하고 있는 모든 사업체이다. 단, 국방 및 국가기밀 보안 관련 시설, 농림어업총조사 대상인 개인이 경영하는 농림어업 사업체, 가구 내 고용 활동 및 달리 분류되지 않은 자가소비 생산 활동, 국제 및 외국기관 등은 제외한다.
경제총조사는 전국의 약 1만 4천 명의 조사요원을 동원하여 2026년 6월 12일부터 7월 22일까지 실시한다. 조사는 원칙적으로 조사원이 사업체를 방문하여 응답자 면접을 통해 조사표를 작성하여야 하나, PC나 모바일 등 온라인, 전화, 팩스, 이메일을 활용한 스마트조사를 추가로 실시할 예정이다.
국가데이터처는 조사계획 수립, 조사표와 조사용품 제작, 통합시스템 구축, 조사요원 교육 등 전반적인 조사 준비를 담당하고, 지방자치단체 특히 시군구는 조사요원 채용 및 관리, 현장조사 등 본조사 실시를 담당한다.
이번 경제총조사의 조사항목은 총 37개 항목으로 사업체의 산업분류와 관계없이 모든 사업체에서 조사하는 공통항목 11개와 사업체의 산업분류에 따라 조사하는 특성항목 26개로 구성되어 있다.

〈 연도별 경제총조사 조사항목 수 〉

(단위: 개)

2010년 기준	2015년 기준	2020년 기준	2025년 기준			
			1차 시험	2차 시험	시범예행	본조사(안)
45공통12, 특성33	39공통16, 특성23	37공통13, 특성24	30공통13, 특성17	38공통13, 특성25	37공통11, 특성26	37공통11, 특성26

〈 2025년 기준 경제총조사 조사항목 〉

공통 항목(11개)	특성 항목(26개)
사업체명, 대표자명, 소재지, 사업자 등록번호, 창설연월, 조직형태, 사업의 종류(산업분류), 영업기간, **인공지능(AI) 활용 여부**, 종사자 수 및 연간 급여액 **(외국인 포함)**, 사업실적	일일 평균 영업시간, **스마트농장운영 여부, 스마트공장운영 여부, 로봇 활용 여부**, 보조금, 무인 결제 기기활용 여부, 온라인 거래 여부, **무인매장 운영 여부**, 사업체 건물(사업장) 연면적, 객실 수, 객석 수, 판매 유형별 매출액 구성비,직능별종사자 수 등

※ 행정자료 대체 항목(유형자산기업, 무형자산기업, 자산총계, 자본금, 자본잉여금, 결산마감월)

2025년 기준 경제총조사에 새롭게 추가된 조사항목은 6개 항목이다.

국가데이터처는 “기술이 주도하는 초혁신경제” 기조에 맞게 ‘인공지능(AI)과 로봇’ 활용 항목을 새롭게 추가하여 인공지능 전환 정책의 기초자료를 제공할 예정이다. 또한 산업 현장의 ‘스마트공장과 스마트농장’ 도입 현황을 파악하여 중소기업 경쟁력 강화와 농업의 디지털 전환을 뒷받침하고자 한다. 더불어 이주민 노동자 지원을 위한 ‘외국인 종사자 수’ 및 비대면 서비스 확산에 따른 국민 안전 등의 기초자료로 ‘무인매장 운영 여부’ 항목도 신규로 조사할 예정이다.

대신 응답자의 부담을 줄여주기 위해 운영장소, 기업체의 유·무형자산 항목은 행정자료로 대체하여 현장조사는 미실시하며, 증기, 냉온수 및 공기조절공급업의 ‘연간 생산량’ 등 특성항목 2개를 폐지하였다. 또한 공통항목 11개 중 사업체명 등 7개 항목은 행정자료를 가지고 조사표에 출력하여 현장조사의 부담을 완화하였다.

2025년 기준 경제총조사 참여

경제총조사는 전체 산업별 규모, 고용, 생산, 비용 등에 관한 정보를 통해 산업·지역 간 구조 비교 및 변화 추이 파악, 조사기준, 조사단위, 조사범위 등을 표준화하여 국가 전체 산업구조에 대한 통일된 통계 기준 제공, 국가·지역 경제정책 수립과 국민소득, 지역소득 등 주요 경제통계의 품질향상을 위한 기초자료 제공을 위해 실시한다.

2025년 기준 경제총조사의 조사대상 사업체는 2026년 6월 12일부터 조사원이 방문하여 참여할 수 있지만, 온라인조사는 6월 1일부터 우선 참여가 가능하며 6월 30일까지 PC나 모바일을 통해 참여할 수 있다.

사업체가 온라인조사 참여를 원한다면 국가데이터처에서 발송하는 온라인조사 안내문에 수록된 온라인조사 참여번호와 초기 비밀번호를 가지고 시간과 장소에 구애 없이 참여가 가능하다. 또한, 응답자의 편의를 위해 구축된 CATI, 웹팩스, 웹메일 등 다양한 방법을 통해 스마트조사 참여가 가능하다.

경제총조사의 성공적 수행을 위해서는 국민들의 적극적인 참여와 정확한 응답이 절실하다. 국가데이터처는 경제총조사의 조사 결과가 국민에게 필요한 정책으로 되돌아갈 수 있도록 최선을 다할 것이며, 2025년 기준 경제총조사에 국민의 적극적인 협조와 많은 관심을 부탁드린다.



AI·빅데이터 시대, 5년 주기 농림어업총조사 '대변신'의 조건

안동환 | 서울대학교 농업생명과학대학 농경제사회학부 교수











전국의 모든 농림어가와 행정리를 대상으로
농림어업 경영 가구와 가구원의 구성, 경영형태와 경영규모 등을 파악하기 위한
농림어업총조사는 대표적인 국가기본통계조사의 하나이다.

조사를 통하여 파악된 정보와 행정리 단위의 소지역 통계는 급변하는 대내외적 환경 변화에 대응한 신속하고 능동적인 농림어업 정책 마련과 농산어촌 지역개발계획을 위한 기초자료로 활용된다. 구체적으로는 농림어가 및 인구, 전 겸업·경영형태 별 가구 구성, 경지면적과 보유어선, 판매 금액 등 농림어가와 농림어업에 대한 중요한 정보가 많은 연구자들에 의해 개인 및 가구단위 또는 지역단위 그리고 국가 간 비교 분석 등에 매우 폭넓게 활용되고 있다. 동시에 농림어업총조사는 농림어가의 모집단을 구축하여 농림어업분야 각종 표본조사의 표본추출틀을 제공함으로써 모든 농림어업 및 농어촌 통계의 근간이 되어 왔다.
하지만 최근 농어촌의 공간구조 변화와 인구구조 변화 등에 따라, 전수조사 및 대면조사를 원칙으로 하는 농림어업총조사도 다음과 같은 환경 변화에 대응할 필요가 있다.

첫째, 전자 조사표와 인터넷 조사 도입으로 고효율·저비용 조사, 정확성 제고 및 자료처리 시간 단축 등을 위해 노력하고 있으나, 급변하는 농림어업 대내외적 여건 변화를 고려할 때 5년 주기의 전수 조사에 따른 자료의 시의성 문제가 꾸준히 제기되어 왔다. 이는 등록센서스로 전환된 인구총조사와 같이 농림어업총조사에도 행정통계기반 센서스 도입이 추진되고 있는 배경의 하나이기도 하다.

둘째, 행정통계 기반 인구총조사 전환에 따른 농림어가 모집단 파악의 어려움이다. 농림어업총조사의 주요 목적 중 하나는 농림어업 부문 각종 표본조사의 모집단과 표본추출틀 제공이다. 하지만 농림어업총조사 모집단 구축의 기반인 인구주택총조사가 등록센서스로 전환됨에 따라 조사 대상 농림어가 모집단 파악이 어려워졌으며, 이에 모집단 구축을 위한 행정자료의 활용이 필요하다.

셋째, 방문 면접을 통한 현장조사 방식의 농림어업총조사는 조사 인력 및 예산 등 조사 부담이 높을 뿐만 아니라 조사 응답의 성실성과 신뢰성에 영향을 줄 수 있다. 특히 고령화 등 농촌인구구조 변화와 도시로부터의 출퇴근 경작농 증가 등으로 인한 부재·불응 가구 증가는 방문면접을 통한 현장조사 방식의 한계를 가져온다. 또한, 농촌인구의 고령화와 1인가구 증가, 경작지와 거주지의 주소 불일치 사례 증가 등에 기인한 응답률 하락으로 현장조사의 어려움이 증가하고 있다.





농림어업과 농어촌의 현황을 파악하고 여건 변화에 따른 정책 대응을 위한 기초자료를 제공하는 주요 국가통계인 농림어업총조사는 농어촌 공간구조와 인구구조 변화와 빅데이터 및 인공지능 활용이 증대되는 흐름에서 새롭게 거듭날 필요가 있다.

먼저 등록센서스 전환에 따른 인구주택총조사의 농림어가 모집단 정보 제공의 한계를 고려한 농림어업총조사의 목적에 대한 재검토와 함께 새로운 모집단 구축 방안 마련이 필요하다.

둘째, 각종 행정 자료와 행정 통계의 연계 확대가 필수적이며, 농림어업총조사를 모집단으로 하여 실시하는 다른 조사(예:농가경제조사, 농림어업조사 등)와의 차이 문제를 최소화하는 동시에 연계 활용할 수 있도록 개선할 필요가 있다. 예를 들면 농림어업 관련 타 통계 조사에 존재하는 항목들을 농림어업총조사 항목과 시계열을 유지하는 방식으로 추가하는 것을 검토할 수 있다. 또한, 국가데이터처로의 승격에 따른 위상에 걸맞게 마스터키(ID) 정보를 통해 다른 조사와 연결성 및 개방성을 높이는 등 관련 부처 간의 협력과 역할 분담을 위한 적극적 노력이 필요하다.

셋째, 농림어업 패러다임 및 농어촌 변화에 대응하여 통계 조사 항목의 추가 및 보완, 조사 항목 정의의 명확화, 조사 대상 단위의 개선 등 통계 수요 대응 노력 강화가 필요하다. 농림어업의 생산 기반에 대한 정확한 통계를 위한 기초자료 제공뿐만 아니라, 곤충, 발효 미생물, 세포 농업 등 새로운 농업 생산의 개념 등장에 따른 농림어업의 범위에 대한 반영이 이루어져야 한다. 또한, 농림어가의 동태적 변화 이해를 위한 패널 분석의 수요, 생활인구와 관계인구 관련 통계 생성, 농촌주민의 삶의 질 관련 통계정보의 생성 등을 고려할 수 있을 것이다.

넷째, 통계정보의 이용자 입장에서 농림어업총조사의 목적, 통계 범위 및 내용에 대한 명확한 홍보, 용어 통일 노력, 그리고 활용도 제고 등이 필요하다. 이를 위해서는 타 기관과의 정보 연계와 정보 접근성 및 정확성 제고, 홍보 강화 등이 필요하다. 개인정보보호와 관련한 제도 개선과 블록체인기술을 활용한 암호화 등에 기반한 연구자의 원시데이터에 대한 접근성 및 활용도 제고가 통계품질 제고를 위한 방안이기 때문이다.

다섯째, 사회관계망서비스(SNS) 데이터, 교통량 데이터, 검색어 데이터 등 다양한 빅데이터와 인공지능 활용 확대에 대응한 통계 방법론의 개발과 활용이다. 이를 위해서는 기존 국가통계 공적 기관들뿐만 아니라 빅데이터를 생산하고 있는 민간 경제주체들과의 교류 협력 방안 마련이 시급하다.

우리나라 농림어업분야의 통계는 다른 나라에 비해 매우 잘 갖추어져 있으며, 특히 농림어업총조사는 국가기본통계로서 오랫동안 정책 및 연구를 위한 기초자료로서 중요하게 활용되어 온 것으로 평가된다. 최근 빅데이터 및 인공지능 활용의 확대와 농림어업부문의 패러다임 변화에 따른 농림어업 정보 통계 수요에 대응하여 농림어업총조사의 발전을 위한 관련 행정부처 및 민간부문과의 협력 확대와 이를 위한 국가데이터처의 위상과 적극적 역할을 기대한다.



경제 전체를 바라보는 통계의 창, 경제총조사의 가치와 미래

장우현 | 한국조세재정연구원 선임연구위원



경제 전체를 바라보는 유일한 창

한 나라의 경제를 왜곡 없이 정확하게 이해하기 위해서는 어떤 통계가 필요할까. 단편적인 추정치나 표본조사만으로는 명확한 한계를 가진다. 모든 경제 주체가 상호 연결된 복잡한 국민경제를 제대로 파악하려면, 경제 전체를 일관된 기준으로 한 시점에서 실사 조사하는 통계가 반드시 필요하다.

이러한 의미에서 전수 경제 센서스 조사는 한 나라의 경제를 이해하는 가장 기본적인 통계 자산이다. 그러나 그만큼 조사와 통계 산출의 난이도도 높기 때문에 형식에 그치지 않고 실질적으로 만족할 만한 경제 센서스를 시행하는 나라는 많지 않다.

다행히 우리나라는 2010년부터 시작된 경제총조사가 그 역할을 충실히 수행하고 있다. 경제총조사는 5년 주기로 실시되는 우리나라 유일의 전수 경제조사로써, 표본이나 추계에 의존하지 않고 실질 조사와 신뢰도 높은 행정자료를 통해 경제의 전체 모습을 일관성 있게 파악한다. 이처럼 전국의 사업체와 기업을 대상으로 하는 대규모 조사를 통해서만 우리는 대한민국 경제의 정확한 실체를 확인할 수 있다.

경제의 큰 그림 1: 부분이 아닌 전체, 생태계를 보다

경제총조사는 경제의 큰 그림이다. 그 첫 번째 특징은 부분이 아닌 전체, 즉 생태계를 본다는 점이다. 국민경제를 연구할 때 특정 부문만 별도로 조사한 통계로는 경제를 온전히 이해할 수 없다.

예를 들어, 중소기업 정책을 연구한다고 가정해 보자. 만약 연구자나 정책 담당자가 중소기업만을 조사한 통계를 근거로 한다면, 전체 경제 속에서 중소기업의 의미와 구조적 역할을 제대로 파악하기 어렵다. 중소기업이 차지하는 경제적 비중, 대기업 및 중견기업과의 관계, 성장 경로와 생태계 내 위치 등은 모두 경제 전반의



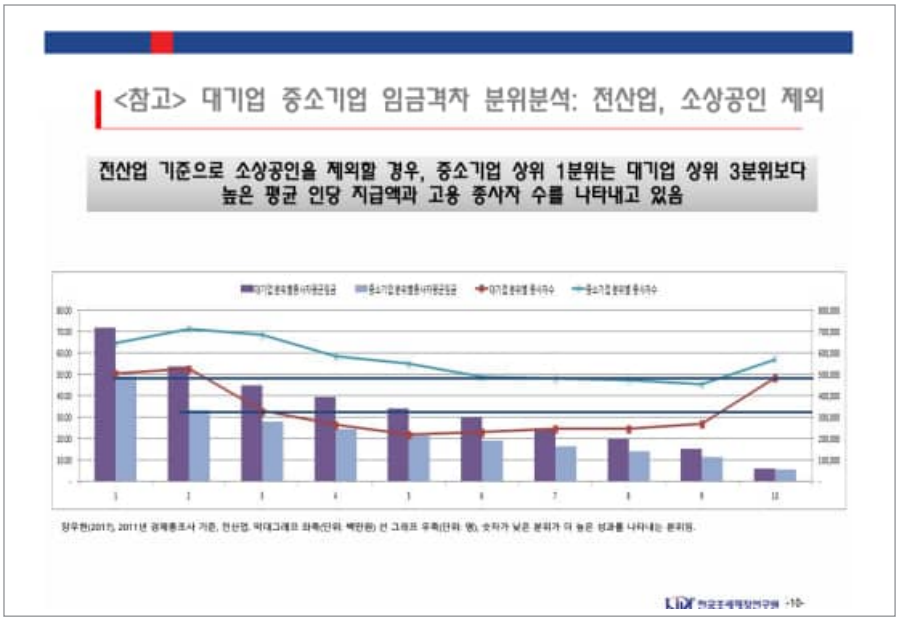
맥락에서만 이해될 수 있다. 경제 생태계 전체를 살피지 않고 중소기업만을 통계 대상으로 살펴본 정치 및 경제 분석은 한계가 있으며 도리어 왜곡과 오해를 야기할 수 있다. 부분이 아닌 전체, 생태계 하에서 개별 주체를 이해해야 하는 대표적 사례이다.

경제의 큰 그림 2: 전체를 보되, 세밀하게 본다

경제총조사의 두 번째 특징은 전체를 보되 동시에 세밀하게 본다는 것이다. 크기가 크다고 해서 해상도가 낮은 그림이 아니다. 오히려 세부 데이터를 정밀하게 담아내는 경제총조사는 부분적으로 세밀하되 경제 전체를 보여주기 때문에, 시스타나 성당의 천장화와 같은 '진정한 큰 그림'이다.

2025년 8월 대통령실 주최 간담회에서 이 같은 경제총조사의 장점이 여실히 드러났다. 기업을 단순히 대기업·중소기업으로 이분화해 평균만 비교하면 오류가 생기지만, 경제총조사의 마이크로데이터는 보다 세밀한 분석을 가능하게 했다. 아래 슬라이드와 같이 중소기업과 대기업을 각각 10개 분위로 구분해 비교한 결과, 상위 중소기업 그룹이 일부 대기업 그룹보다 인당 지급액과 고용 규모에서 더 우수한 성과를 보인다는 사실을 확인할 수 있다.

즉, 모든 대기업이 우월하고 모든 중소기업이 열등하다는 통념은 경제총조사 데이터로 명확히 반증 가능하다. 이는 경제총조사가 우리 경제의 실상을 근거 기반으로 보여주는 '정확한 큰 그림'임을 의미한다.



모든 통계의 기준점, 도량형의 역할

경제총조사는 다른 통계들의 기준점, 즉 도량형 역할을 한다. 이는 통계학적으로 매우 중요하나 많은 이들이 간과하는 부분이다.

GDP와 같은 주요 거시 통계 역시 경제총조사와 연계되어야 정확도와 일관성을 확보할 수 있다. 실제로 2024년 한국은행이 국민계정의 기준년을 2020년으로 개편할 때, 경제총조사 자료를 핵심적으로 반영하였다. 경제총조사와 실측 투입산출표 등의 자료가 결합되면서 우리나라 경제 규모가 보다 정밀하게 측정될 수 있었다.

이처럼 경제총조사는 통계체계의 중심을 잡고 일관성을 유지하게 하는 기준점으로, 다른 어떤 조사로도 대체 불가능한 존재다. 경제총조사를 다른 통계의 대체자료로 오해하는 사람들이 있다. 다른 통계들이 있는데 굳이 경제총조사가 필요한가라는 질문을 던지는 사람들이다. 이는 대체성과 보완성에 대한 오해에 의한 것으로 볼 수 있다. 경제총조사는 다른 통계로 대체될 수 있는 통계가 아니고, 다른 통계들의 든든한 보완 자료로서 활용된다는 사실을 모두 이해할 필요가 있을 것이다.

연구와 정책 의사결정의 첫 페이지

이러한 이유로 필자는 국제기구 회의나 정부 보고를 포함한 주요 연구와 정책 발표를 진행할 때마다 가장 첫 페이지에 항상 경제총조사에 기반한 요약 자료를 제시할 수밖에 없었다. OECD 회계 관련 국제회의와 같은 자리에서 경제총조사가 제공하는 사업체 수·종사자 수·매출액 등의 기초통계는 항상 경제·재정 분석의 출발점이 된다.

경제총조사가 연구와 정책 의사결정의 기초 측면에 기여하는 한 예로, 필자가 2013년 한국개발연구원(KDI)에서 수행한 「중소기업 지원정책 개선 연구」에서 최초로 경제총조사를 활용한 사례를 소개하고자 한다. 조사 당시 경제총조사는 중소기업을 기업이 아닌 사업체 단위로만 보던 오류를 바로잡고, 정책당국이 보다 정확한 중소기업 통계를 구축하는 계기를 마련했다. 그 결과 과거 대기업의 소규모 사업체를 모두 중소기업으로 분류하던 왜곡을 교정하는 근거자료로 활용되었다.

참여가 만드는 정확한 통계

무엇보다 경제총조사의 성공은 국민과 사업체의 적극적인 참여에 달려 있다. 물론 모든 국민에게 통계법에 따른 법적 응답 의무가 있지만, 그보다 더 중요한 것은 이 조사가 우리 모두의 이익을 위한 공동 자산이라는 인식이다.

경제총조사 결과는 앞선 예시들에서도 구체적으로 확인할 수 있듯이 국가 산업 구조와 변화를 파악하여 각종 경제·산업정책의 기초자료로 활용된다. 또한, 사업체 표본조사들의 모집단 기준이 되며, 시장집중도와 산업다각화지수 등 주요 통계산출의 토대를 제공한다. 나아가 대학·연구기관·민간 기업의 연구와 경영 분석에도 폭넓게 쓰인다.

한 사람, 한 사업체의 참여가 곧 대한민국 경제정책의 정확성을 높인다. 그 혜택은 결국 우리 모두에게 돌아온다. 2025년 경제총조사를 앞둔 지금, 각 사업체의 성실한 참여를 진심으로 당부드린다.

경제총조사의 미래: 더 자주, 더 정확하게

이처럼 중요한 경제총조사를 5년마다 한 번만 실시해 아쉬움이 크다. 변화 속도가 빨라진 현실을 감안할 때, 2년 주기 조사로의 전환은 충분히 검토할 가치가 있다. 국가데이터처 장래인구추계를 5년에서 2년 주기로 단축했던 것처럼, 경제총조사 주기 단축은 경제정책의 시의성과 대응력을 높이는 효과를 가져올 수 있다.

물론 예산과 인력의 제약이 있지만, 디지털 기술과 행정자료의 적극적인 활용을 통해 충분히 실현 가능한 목표라고 본다. 논증을 위해 경제총조사를 다시 인용하자면, 우리나라 경제는 2020년 기준 매출액 6,711조원 규모의 거대한 시스템이다. 그 규모를 고려하면 경제총조사에 투입되는 예산은 극히 미미한 수준이라고도 볼 수 있을 것이며, 조사로 얻는 편익은 그 비용을 훨씬 상회할 가능성이 높다.

따라서 필자는 비록 재정 절약을 항상 염두에 두어야 하는 재정전문가임에도, 경제총조사에 있어서만큼은 더 자주, 더 정밀한 조사를 통해 우리나라 경제정책의 정확성과 대응력을 높여 나가야 한다고 생각한다.

경제총조사의 미래를 응원하며

2010년 첫발을 내디딘 이후, 경제총조사는 대한민국 경제의 나침반 역할을 충실히 수행해 왔다. 전국 사업체의 현황을 빠짐없이 조사하여 경제 구조를 정확히 보여주고, 모든 경제통계의 기준점을 마련했다.

앞으로도 경제총조사가 더욱 발전하여 우리 경제정책의 든든한 토대가 되기를 기대한다. 국가데이터처로 승격한 통계청을 비롯한 관계기관의 노고에 깊이 감사드리며, 외부 연구기관의 일원으로서 그 발전을 지속적으로 지원하고 응원하고자 한다.

정확한 데이터가 정확한 정책을 만들고, 정확한 정책이 국민의 삶을 개선한다는 믿음으로, 경제총조사의 성공적인 수행을 진심으로 기원한다.



데이터로 배우는 시대, 실용 통계와 소프트웨어 교육의 만남

김갑수 | 서울교육대학교 정보교육과 교수



1 서론

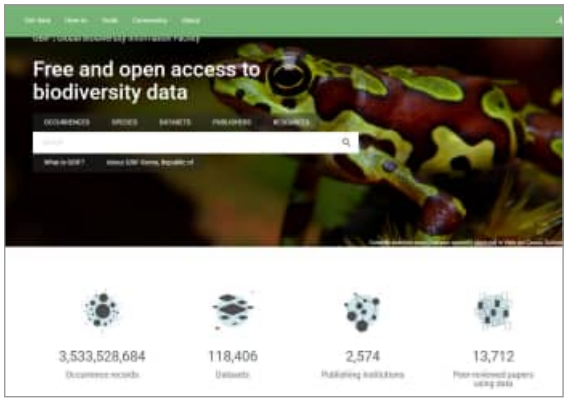
현재 2022개정 교육과정을 기반으로 구성된 교과서를 바탕으로 초등학교, 중학교, 고등학교에서는 교육을 진행하고 있다. 2022개정 교육과정에서는 언어 능력, 수리 능력 및 디지털 능력을 매우 중요한 능력으로 강조하고 있다. 디지털 능력은 실용 통계와 소프트웨어 교육의 핵심 요소가 된다. 지금까지 대부분 교과의 교육 내용을 구성할 때 통계 자료는 교과서를 작성하는 시점의 지난 데이터 또는 가상의 데이터들을 기반으로 편성되었다. 또한, 디지털 능력의 경우에는 단순한 도구, 즉 소프트웨어의 사용에만 초점을 두고 있었다. 현재는 살아있는 교육, 즉 학교 교육이 사회와 연계되기 위해서는 실질적인 실용 통계와 소프트웨어 교육이 필수적이다. 이를 기반으로 인공지능 시대에 각 영역에서 주도적인 삶을 살아내는 방법을 교육할 수 있을 것이다.

2 실용 통계

이제 우리는 실시간 데이터들을 국가통계포털(https://kosis.kr/), 공공데이터포털(https://www.data.go.kr/), 기상자료개방포털(https://data.kma.go.kr/cmmn/main.do) 및 유럽데이터포털(https://data.europa.eu/)등에서 찾아볼 수 있다. 네이버의 데이터랩(https://datalab.naver.com/)에 가보면 실시간 검색 데이터들을 볼 수 있고, 구글트렌드(https://trends.google.com/trends/)에 가 보면 전 세계 검색 트렌드들을 실시간으로 과거와 비교해 볼 수 있다. 플라이트어웨어(https://www.flightaware.com/live/) 에 가보면 현시점에서 비행기들이 지구의 어느 위치에 있는지 알려준다. 인공위성으로 구름 모양을 직접 볼 수 있고, 전 세계 CCTV로 직접 교통상황을 볼 수 있다.



[그림 1] AI 허브



[그림 2] 국제적 생물다양성 정보 플랫폼



[그림 3] 클라우드 기반 지리공간 데이터 분석 플랫폼

한국지능정보사회진흥원(https://www.nia.or.kr/site/nia_kor/main.do)에서는 국가적으로 구축해 둔 인공지능 학습 데이터 집합(https://www.aihub.or.kr/)이 있고 구체적인 내용은 다음 [그림1]에 제시되어 있다. 간단히 살펴보면 2025년 10월 22일 현재 907건의 데이터가 구축되어 있다.

GBIF(https://www.gbif.org/)에서는 국제적인 생물다양성 플랫폼 [그림2]을 제공하며, 생태학적 연구 및 환경 보호에 중요한 데이터를 제공한다. 생물의 데이터 집합이 2025년 10월 22일 현재 118,406건이 있고, 이 데이터를 이용한 연구 논문들이 13,712개 있다.

구글 지구엔진(https://earthengine.google.com/)에서는 위성영상과 지구 관측 데이터를 제공한다. 이 데이터들은 30년 이상의 역사적 이미지와 과학 데이터 세트가 포함되어 있으며, 매일 업데이트 및 확장되고 즉시 분석할 수 있는 80페타바이트(Petabyte, PB) 이상의 지리 공간 데이터가 포함되어 있다.

이처럼 이제는 우리 생활 속에 발생한 다양한 데이터들을 이용할 수 있는 환경이 되었다. 따라서 2022개정 교육과정의 각 교과에서 내용을 구성할 때 위와 같은 데이터를 활용한 실용 통계 교육을 진행할 수 있다. 가상의 데이터를 활용하기보다 현재 실시간으로 변화하는 데이터를 즉각 활용해 문제를 해결하는 교육이 필요하다.

3 교육과정과 실용 통계

2022개정 교육과정에서 초등학교 수학 교과에 활용할 수 있는 실용 통계를 간단히 보자. 초등학교에서 배워야 하는 '수학의 자료와 가능성'이라는 큰 분야가 있다.

[표1]의 각 성취 기준에 이용할 자료들은 앞에서 논의한 실시간 데이터들이다. 예를 들어 국가데이터처 자료나 네이버 데이터랩 등의 자료를 이용하여 위의 성취 기준별로 데이터 집합을 학생들이 직접 만들어보는 것이다. 값이 많이 축적된 데이터를 다양하게 분석하고 정리하기 위해서 소프트웨어가 필수적으로 사용된다. 초등학교 저학년들은 소프트웨어를 이용하는 방법과 체험하는 것에 초점을 맞추고, 초등학교 고학년의 경우에는 직접 간단한 소프트웨어들을 개발해 보는 것이 좋다.

[표 1] 2022개정 교육과정 수학과 교육과정(교육부, 2022년)

[2수04-01] 여러 가지 사물을 정해진 기준 또는 자신이 정한 기준으로 분류하여 개수를 세어 보고, 기준에 따른 결과를 말할 수 있다.

[2수04-02] 자료를 분류하여 표로 나타내고, 자료를 표로 나타내면 편리한 점을 말할 수 있다.

[2수04-03] 자료를 분류하여 ○, ×, / 등을 이용한 그래프로 나타내고, 자료를 그래프로 나타내면 편리한 점을 말할 수 있다.

[4수04-01] 자료를 수집하여 그림그래프나 막대그래프로 나타내고 해석할 수 있다.

[4수04-02] 자료를 수집하여 꺾은선그래프로 나타내고 해석할 수 있다.

[4수04-03] 탐구 문제를 해결하기 위해 자료를 수집, 정리하여 막대그래프나 꺾은선그래프로 나타내고 해석할 수 있다.

[6수04-01] 평균의 의미를 알고, 자료를 수집하여 평균을 구하고 해석할 수 있다.

[6수04-02] 자료를 수집하여 띠그래프나 원그래프로 나타내고 해석할 수 있다.

[6수04-03] 탐구 문제를 설정하고, 그에 맞는 자료를 수집, 정리하여 적절한 그래프로 나타내고 해석할 수 있다.

[6수04-04] 사건이 일어날 가능성을 말로 표현하고 비교할 수 있다.

[6수04-05] 사건이 일어날 가능성을 수로 나타낼 수 있다.

[6수04-06] 자료를 이용하여 가능성을 예상하고, 가능성에 근거하여 적절한 판단을 내릴 수 있다.



2022개정 교육과정에서 중학교 과학 교과에 활용할 수 있는 실용 통계를 간단히 보자.

[표 2] 2022개정 교육과정 수학과 교육과정(교육부, 2022년)

[9과17-01] 지구 대기권을 4개 권역으로 구분하며, 온실효과와 지구온난화를 복사 평형의 관점으로 설명할 수 있다.
[9과17-02] 대기 대순환에서 위도별 바람의 특성을 파악하고, 대기 대순환의 역할을 설명할 수 있다.
[9과17-03] 상대습도, 단열 팽창 및 응결 현상의 관계를 이해하고, 구름의 생성과 강수 과정을 설명할 수 있다.
[9과17-04] 기압, 기단, 전선의 개념을 이해하고, 일기도에서 저기압과 고기압의 분포에 따른 날씨를 해석할 수 있다.

〈탐구 활동〉

- 기후변화 관련 자료 검색 및 분석하기
- 디지털 탐구 도구를 이용하여 구름 생성 과정 실험하기
- 일기도, 레이더 및 위성 영상 등 실시간 데이터를 활용하여 날씨 해석하기

앞서 언급한 실시간 데이터를 이용해 표2의 각 성취 기준에 맞춰 자료를 활용할 수 있다. 기상청 자료나 다른 기상정보들의 자료를 이용하여 위의 성취 기준별로 데이터 집합을 만들어서 학생들이 직접해 보는 것이다. 실제 구름 생성 과정을 시뮬레이션하는 매우 간단한 소프트웨어도 직접 만들어 볼 수 있으면 좋겠다.

4 소프트웨어 교육

우리나라 소프트웨어 교육은 정책 및 국가적인 강조에 비해 실질적으로 매우 적은 시수로 운영되고 있다. 2022개정 교육과정에서 초등학교 6년 동안 34시간의 교육을 권장하고 있으나 제대로 이행되지 못하고 있다. 초등학교 교과목에 '정보'는 포함되지 않으며 소프트웨어 교육은 실과 과목의 한 영역으로 포함되어 있다. 초등학교 6학년이 되어서야 비로소 해당 교육을 받을 수 있다. 따라서 학생들의 소프트웨어 교육 수준의 격차가 크다. 다음 표3은 초등학교 실과 교과의 소프트웨어 영역 교육과정이다.

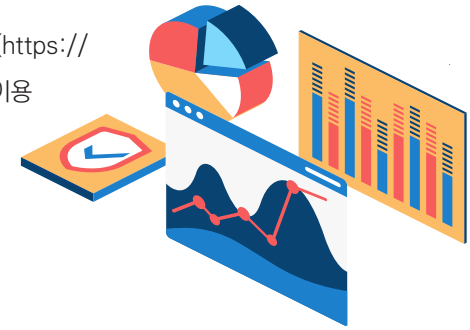
[표 3] 2022개정 교육과정 실과정보(교육부, 2022년)

[6실05-01] 컴퓨터를 활용한 생활 속 문제해결 사례를 탐색하고 일상생활 속 문제를 해결하기 위한 알고리즘을 다양한 방법으로 표현한다.
[6실05-02] 컴퓨터에게 명령하는 방법을 체험하고, 주어진 문제를 해결하는 프로그램을 작성한다.
[6실05-03] 실생활의 문제를 해결하는 프로그램을 협력하여 작성하고, 산출물을 타인과 공유한다.
[6실05-04] 디지털 데이터와 아날로그 데이터의 특징을 이해하고, 인공지능에 활용할 수 있는 데이터의 유형이나 형태를 탐색한다.
[6실05-05] 인공지능이 만들어지는 과정을 체험하고, 인공지능이 사회에 미치는 영향을 탐색한다.

소프트웨어 교육의 교과 시간은 매우 적기 때문에 타 교과 시간에서 실용 통계를 기반한 소프트웨어 교육을 병행하는 것은 매우 중요하다.

각 교과의 내용을 학습하는 과정에서 실용 통계 교육을 진행하기 위해 소프트웨어를 선택하고 데이터를 분석하는 건 필수적이다. 소프트웨어를 선택할 때는 에듀테크, 인공지능 기업들이 만든 인공지능 소프트웨어를 이용하거나, 직접 소프트웨어를 만들어 사용하는 방법이 있다. 직접 소프트웨어를 만드는 방법은 간단한 라이브러리를 사용할 수 있게 만들어 둔 모듈을 이용하는 방법, 인공지능을 이용해 소프트웨어를 만드는 방법, 마지막으로 파이썬, C, 자바 등의 툴을 이용해 직접 만들어 사용하는 방법이 있다.

초등학생들의 블록 코딩 및 게임 기반 학습을 위하여 엔트리(https://playentry.org/), Scratch (https://scratch.mit.edu)등을 이용하면 좋고, 중학생들은 파이썬 언어를 이용하기 좋다. 초중등 학생이 소프트웨어 교육에 사용하기 좋은 도구들은 다음 표4와 같다.



[표 4]

이름	URL	주요 기능
Orange	https://orange. biolab.si https://orangedata mining.com/	데이터 마이닝 및 머신러닝 도구, 시각화, 분류, 회귀 분석 제공
Weka	https://www.weka.io	머신러닝 및 데이터 마이닝 소프트웨어, 분류, 클러스터링, 회귀 분석 제공
Grafana	https://grafana.com	실시간 데이터 모니터링 및 시각화 도구, 다양한 데이터 소스와 통합 가능
Bokeh	https://bokeh.org	웹 기반의 대화형 데이터 시각화 도구, Python 라이브러리
Pandas Profiling	https://pandas-profiling.github.io	데이터 분석 리포트 자동 생성, 탐색적 데이터 분석(EDA) 도구
Datawrapper	https://www.datawrapper.de	웹 기반 데이터 시각화 도구, 차트 및 지도 시각화
TensorFlow	https://www.tensorflow.org	머신러닝 및 딥러닝을 위한 오픈소스 라이브러리, Python 및 JavaScript 지원
Vega-Lite	https://vega.github.io	간단한 인터랙티브 시각화 도구, JSON 기반 시각화 언어
엔트리	https://playentry.org/	블록 코딩을 통해 프로그래밍을 학습하고, 데이터 분석 및 인공지능 프로그램 가능

5 결론

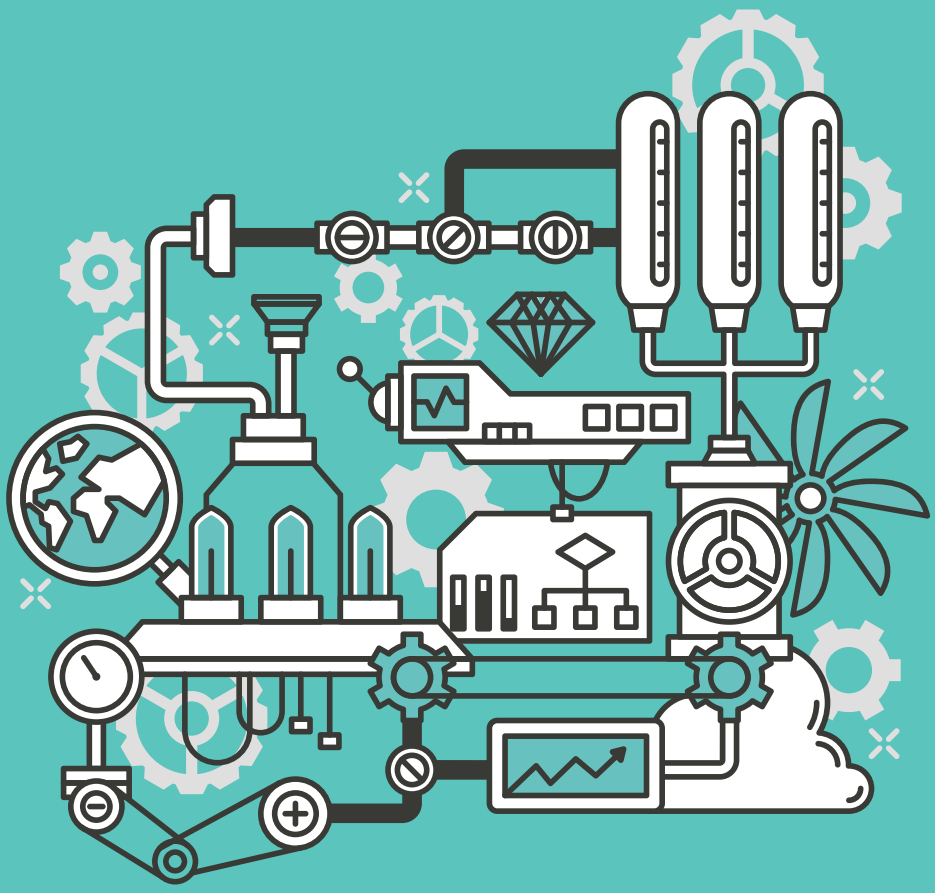
현대 사회는 데이터 중심 사회이다. 데이터 중심은 곧 실용 통계를 말한다. 중요한 정책을 결정하고 과학 연구를 진행할 때, 그리고 삶에서 내리는 순간순간의 판단에는 통계적 사고와 데이터 해석 능력이 필요하다. 이를 위해 디지털 전환 사회에서는 데이터를 읽고, 분석하고, 해석하는 능력인 새로운 문해력이 요구된다. 학생들은 일상 속에서 기온 변화, SNS 여론조사, 유튜브 조회수 등 다양한 형태의 데이터를 접하지만, 이를 체계적으로 해석하고 의사결정에 활용하는 역량은 부족하다. 이같은 역량을 기르기 위해서 초중등 학교의 교육 과정 안에서 실용통계를 기반으로 한 소프트웨어 교육이 필요하다. 소프트웨어 역량이 있으면 데이터를 바로 수집하고 처리할 수 있으며, 다양한 분석 도구들을 이용할 수 있다. 모든 것이 소프트웨어로 되어 있는 시대이다. 실용 통계와 소프트웨어 교육은 단순 코딩을 넘어서 사고력과 문제 해결 능력을 기르는 교육이다.



이름	URL	주요 기능
Scratch	https://scratch.mit. edu	블록 코딩을 통해 프로그래밍 및 AI 개념을 배우는 도구, AI 확장으로 이미지 인식, 음성 인식 등 추가 가능
Teachable Machine	https://teachablemachine.withgoogle.com	이미지, 소리, 포즈 인식 AI 모델을 쉽게 만들 수 있는 도구, 직관적인 인터페이스 제공
Google Colab	https://colab.research.google.com	Python과 AI 모델을 실험할 수 있는 클라우드 기반 환경, 초등학생도 접근 가능
RunwayML	https://runwayml.com	창의적인 AI 프로젝트 도구, 이미지 생성 및 편집, 영상 변환 등 다양한 창의적 작업 지원
AI Experiments	https://experiments.withgoogle.com/collection/ai	AI 기술을 재미있는 실험으로 경험할 수 있는 웹사이트, AI의 다양한 활용 사례를 체험
Cognimates	https://cognimates.me	코딩과 AI를 결합한 학습 플랫폼, 로봇 제어, 이미지 및 음성 인식, 자연어 처리 기능 제공
Kaggle	https://www.kaggle.com	데이터 과학 및 머신러닝 실습, 대규모 데이터셋 활용, AI 및 데이터 분석 대회 참여
Microsoft Azure Machine Learning Studio	https://studio.azureml.com	머신러닝 모델 구축 및 배포, 시각적 인터페이스 제공, 프로그래밍 없이도 AI 실습 가능
Pytorch	https://pytorch.org	딥러닝 및 인공지능을 위한 Python 기반의 오픈소스 라이브러리, 머신러닝 모델 훈련 가능

산업 구조의 재편, 제조업이 그려가는 새로운 5년

전상범 사무관, 김근일 주무관 | 국가데이터처 산업통계과



01 들어가며...

지난 5년(2019~2023년), 우리 경제를 둘러싼 환경 변화는 국내 제조업에 큰 영향을 남겼다. 본 분석은 이러한 변화를 짚어보고, 앞으로의 지속 가능한 성장 전략과 정책 수립에 필요한 시사점을 제공하고자 한다.

제조업 구조 변화 분석을 위해, 국가데이터처 「광업·제조업 조사」 원자료를 활용하여 제10차 KSIC 중분류 25개 업종을 분석하였다. 자동차·석유화학·철강·반도체·조선 등 핵심 5대 업종은 소분류 업종별로 동향을 추적하였다.

출하액, 부가가치액, 사업체 수, 종사자 수, 급여액, 주요 생산비, 유형자산 연말잔액의 주요 7개 지표를 분석하였다. 주요 지표를 활용하여 산업 활력도, 비용 구조, 생산성, 경영 성과 등의 구조 분석 프레임으로 변화를 살펴보았다.

02 제조업의 구조적 변화를 살펴보자!

1. 코로나19에도 강한 대한민국 제조업 '활력 지표'

코로나19 시기에도 우리 제조업은 흔들리지 않았다. 출하액은 연평균 6.6% 늘었고, 유형자산은 5.5%, 사업체 수는 1.2% 증가하며 충격을 빠르게 회복하였다. 반도체·바이오·친환경 소재 같은 신산업 성장과 수출 반등이 이러한 회복을 이끌었다.

〈표 II-1〉 최근 5년간 제조업의 활력도 추이

단위: %

	2019	2020	2021	2022	2023	연평균 증감률
사업체 수 (천 개)	69.6	70.8	72.5	73.3	73.1	1.2
종사자 수 (천 명)	2,928.3	2,910.4	2,937.9	2,981.8	2,974.3	0.4
사업체당 평균 종사자 수 (명)	42.0	41.1	40.5	40.7	40.7	-0.8
유형자산 연말잔액 (조원)	617.1	643.9	695.7	715.8	765.5	5.5
출하액 (조원)	1,541.8	1,504.3	1,765.7	2,040.3	1,988.9	6.6

자료: 국가데이터처, 광업제조업조사

제조업의 비용 및 가치 창출 구조 분석에서는 주요 생산비(원재료비, 연료비, 전력비, 용수비, 외주가공비, 수선비)가 7.4% 증가하여 비용 측면에서 위험 요인이 감지되었다. 특히, 국제 연료비 급등, 전력시장 구조 변화, 탄소중립 이행에 따른 탄소 비용 전가 등 복합적 요인으로 증가하였다. 최저임금 인상으로 인하여 급여액 또한 4.2% 증가하였다.

〈표 II-2〉 최근 5년간 제조업의 비용 구조 추이

단위: %

	2019	2020	2021	2022	2023	연평균 증감률(%)
부가가치(조원)	557.0	554.5	641.3	722.3	672.6	4.8
급여액(조원)	135	139	148	155	160	4.2
인당 급여액(백만원)	46.3	47.6	50.4	52.0	53.8	3.8
주요 생산비(조원)	996	948	1,140	1,352	1,323	7.4
주요 생산비/출하(%)	64.6	63.0	64.6	66.2	66.5	-

자료: 국가데이터처, 광업제조업조사

2. 경영 효율성과 산업 경쟁력은 어떻게 변했을까?

자동화와 디지털 전환 덕분에 노동생산성은 연평균 4.4% 증가하였다. 반면, 자본생산성은 0.7% 감소하였다. 설비 투자가 크게 늘었지만, 부가가치 증가 속도가 이를 따라가지 못한 결과이다. 이는 자동화 초기의 과도기적 현상으로 볼 수 있으며, 저생산성 업종 퇴출과 고부가 업종 중심의 재편을 촉진할 수 있다. 다만 자본 효율성 저하는 기업들의 신규 투자 의지를 약화시킬 위험이 있다.

〈표 II-3〉 최근 5년간 제조업의 생산성 추이

단위: %

	2019	2020	2021	2022	2023	연평균 증감률(%)
노동생산성(백만원/인)	190.2	190.5	218.3	242.2	226.1	4.4
자본생산성	0.9	0.9	0.9	1.0	0.9	-0.7
노동장비율(백만원/인)	210.7	221.2	236.8	240.1	257.4	5.1

자료: 국가데이터처, 광업제조업조사

주: 노동생산성=부가가치/종사자 수, 자본생산성=부가가치/유형고정자산, 노동장비율=유형자산 연말잔액/종사자 수

최근 5년간 제조업 생산액은 6.5% 늘었지만, 부가가치는 4.8% 증가에 그쳤다. 그 결과 부가가치율은 2019년 35.9%에서 2023년 33.7%로 떨어지며 경쟁력이 약화되었다. 이는 원자재·에너지·물류비 등 주요 비용이 급격히 상승하면서 원가 부담이 커졌기 때문이다.

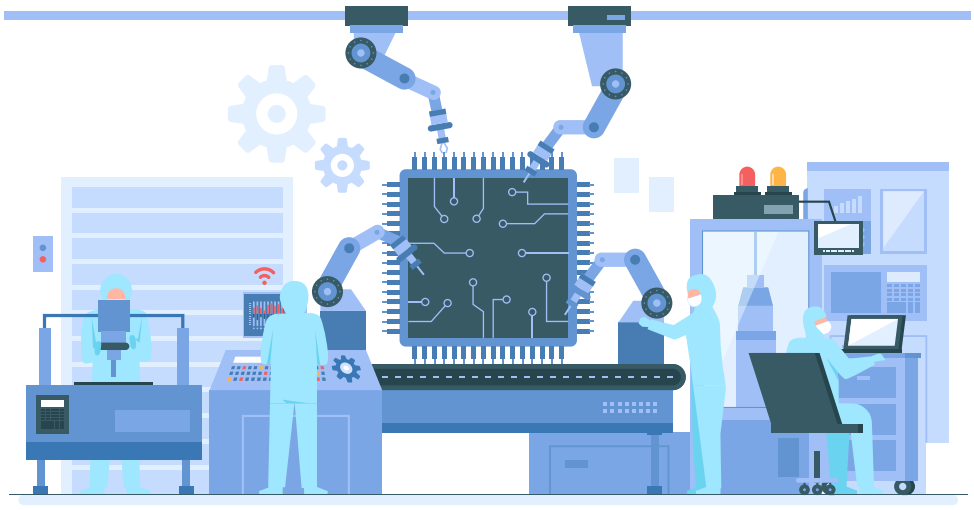
〈표 II-4〉 최근 5년간 제조업의 성과 분석

단위: 조 원, %

	2019	2020	2021	2022	2023	연평균 증감률(%)
생산(A)	1,552.7	1,502.5	1,781.3	2,073.8	1,995.9	6.5
부가가치(B)	557.0	554.5	641.3	722.3	672.6	4.8
부가가치율(B/A)	35.9	36.9	36.0	34.8	33.7	

자료: 국가데이터처, 광업제조업조사, 10인 이상 사업체 기준

주: 부가가치율=(부가가치/생산)*100



03 제조업종별 구조적 변화를 살펴보자!

1. 업종별 성장과 쇠퇴 추이

업종별로 보면 반도체·의약품·의료·정밀·광학 같은 고위기술 산업이 빠르게 부상했다. 사업체 수는 반도체 7.4%, 의약품은 4.7%, 식료품은 4.4%, 종사자 수는 의약품 6.2%, 반도체는 3.2%, 식료품은 3.2%로 증가하였다. 이들 산업은 사업체·고용·투자율이 증가하여 제조업 내 비중이 확대되었다. 반면 섬유·의복·가죽·목재 등 전통 소비재 산업은 중국 등 신흥국과의 가격 경쟁에서 밀리며 비중이 위축됐다.

유형자산 연말잔액은 반도체 15.1%, 담배 14.1% 증가하여 투자가 급격히 확대되었다. 반도체와 전자제품은 글로벌 데이터센터 AI·5G 수요처 확대로 삼성·SK하이닉스 등이 중심이 되어 신규라인 증설과 첨단 공정 등을 도입하여 생산능력이 크게 확장되었다. 담배 제조업은 세계적인 연초 흡연율 감소 속에서 전자담배 시장에 자원을 집중하여 성장을 견인하였다.

〈표 Ⅲ-1〉 주요 제조업종별 사업체 수, 종사자 수, 유형자산 증감률

단위: %

제조업종별 사업체 수 연평균 증감률			제조업종별 종사자 수 연평균 증감률			제조업종별 유형자산 연평균 증감률		
1	(반도체)	7.4	1	의약품	6.2	1	(반도체)	15.1
2	의약품	4.7	2	(반도체)	3.2	2	담배	14.1
3	식료품	4.4	3	식료품	3.2	3	전자제품	10.0
				~ · ~ · ~				
	제조업	1.2		제조업	0.4		제조업	5.5

자료: 국가데이터처, 광업제조업조사

주: 10인 이상 사업체, KSIC 중분류 산업 기준, ( ) 안은 KSIC 소분류 산업 기준

2. 제조업종별 '원가구조 들여다보기'

주요 생산비 증가율도 업종별로 차별화되었다. 특히, 전기장비 14.4%, 반도체 12.5%, 자동차 10.2% 등에서 주요 생산비가 빠르게 증가하며 비용 부담이 커졌다. 이는 반도체, 자동차의 원부자재 가격 변동에 따른 증가 요인에서 기인하였다.

〈표 Ⅲ-2〉 주요 제조업종별 주요 생산비 및 증감률

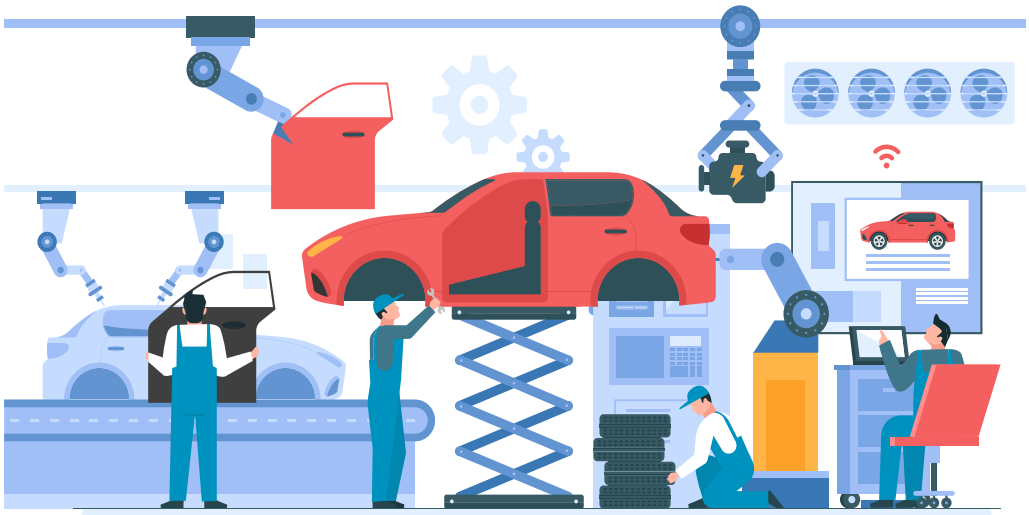
단위: 10억 원, %

2019			2023			연평균 증감률		
1	자동차	141,320	1	자동차	208,714	1	전기장비	14.4
2	전자제품	123,947	2	전자제품	156,724	2	(반도체)	12.5
3	1차 금속	110,314	3	1차 금속	144,797	3	자동차	10.2
				~ ~ ~				
	제조업	995,692		제조업	1,323,392		제조업	7.4

자료: 국가데이터처, 광업제조업조사

주: 10인 이상 사업체, KSIC 중분류 산업 기준, ( ) 안은 KSIC 소분류 산업 기준

주요 생산비가 출하액에서 차지하는 비율은 산업의 수익성과 직결되며, 2023년 제조업 평균 주요 생산비 비중이 2019년 대비 2.0%p 증가하여 수익성이 악화되었다. 2023년 기준으로 1차 금속(78.0%), 조선(76.5%), 철강(75.5%) 등 주력 산업의 주요 생산비 비중이 상대적으로 높은 수준을 유지하고 있다. 주요 생산비 비중이 크게 증가한 업종은 반도체(15.2%p), 전자제품(9.4%p), 조선(7.8%p) 등으로 최근 5년간 수익성이 감소하였다.



〈표 Ⅲ-3〉 주요 제조업종별 주요 생산비/출하액 비중 및 변화

단위: %, %p

2019			2023			증감		
1	석유정제	80.8	1	1차 금속	78.0	1	(반도체)	15.2
2	1차 금속	77.9	2	(조선)	76.5	2	전자제품	9.4
3	(철강)	76.4	3	(철강)	75.5	3	(조선)	7.8
				~ · ~ · ~				
	제조업	64.6		제조업	66.5		제조업	2.0

자료: 국가데이터처, 광업제조업조사

주: 10인 이상 사업체, KSIC 중분류 산업 기준, ( ) 안은 KSIC 소분류 산업 기준

1인당 급여액은 석유정제, 담배, 반도체 순으로 대기업형 산업에서 높은 경향을 보인다. 음료, 가죽·신발, 조선, 업종은 급여 수준은 낮으나, 급격한 고용 감소로 인해 소수 인원에게 집중되어 평균 인건비가 상승하였고, 이에 따라 1인당 연평균 증감률이 큰 폭으로 증가하는 추세를 나타내고 있다.

〈표 Ⅲ-4〉 주요 제조업종별 1인당 급여액 및 증감률

단위: 백만 원, %

2019			2023			연평균 증감률		
1	석유정제	91.8	1	석유정제	109.5	1	음료	6.2
2	(반도체)	76.5	2	담배	79.3	2	가죽·신발	5.9
3	(석유화학)	66.7	3	(반도체)	76.8	3	(조선)	5.4
				~ · ~ · ~				
	제조업	46.3		제조업	53.8		제조업	3.8

자료: 국가데이터처, 광업제조업조사

주: 10인 이상 사업체, KSIC 중분류 산업 기준, ( ) 안은 KSIC 소분류 산업 기준

출하액 대비 인건비 비중이 증가한 업종은 가죽·신발(1.5%p), 의료·정밀·광학(0.8%p), 전자제품(0.7%p) 등으로 인당 인건비가 낮은 업종에서 증가 폭이 크게 확대되어 노동비용 부담이 증가하였다.

〈표 Ⅲ-5〉 주요 제조업종별 급여액/출하액 비중 및 변화

단위: %, %p

2019			2023			증감		
1	인쇄·기록매체	21.3	1	인쇄·기록매체	19.7	1	가죽·신발	1.5
2	기타 제품	16.4	2	기타 제품	16.8	2	의료·정밀·광학	0.8
3	섬유제품	14.6	3	섬유제품	14.8	3	전자제품	0.7
				~ · ~ · ~				
	제조업	8.8		제조업	8.0		제조업	-0.7

자료: 국가데이터처, 광업제조업조사

주: 10인 이상 사업체, KSIC 중분류 산업 기준, ( ) 안은 KSIC 소분류 산업 기준

3. 제조업종별 노동·자본 업종별 생산성 비교

제조업종의 노동생산성은 연간 4.4%로, 설비 및 자동화 투자 확대, 기술혁신 및 공정 개선, 인력 숙련도 향상, 고부가가치 제품 생산 전환, 디지털화 및 업무 혁신 등에 힘입어 꾸준히 증가하였다. 제조업종별 노동생산성 향상은 기술·설비 투자와 구조조정에 따른 증가로 업종별 특성에 따라 상이하다.

석유정제산업은 대기업 중심 과점 구조로 설비 효율성 및 공정 최적화가 고도화되어 노동생산성이 증가하였고, 의복·모피 등은 인력 감소로 인한 생산의 아웃소싱 확대 등 구조조정을 통해 노동생산성이 증가하는 경향을 보였다.

〈표 Ⅲ-6〉 주요 제조업종별 노동생산성 및 증감률

단위: 백만 원/인, %

2019			2023			연평균 증감률		
1	석유정제	1,973.1	1	석유정제	3,553.1	1	석유정제	15.8
2	담배	845.6	2	담배	902.1	2	전기장비	11.3
3	(반도체)	674.1	3	(석유화학)	523.6	3	의복·모피	11.1
				~ · ~ · ~				
	제조업	190.2		제조업	226.1		제조업	4.4

자료: 국가데이터처, 광업제조업조사

주: 1) 노동생산성=부가가치/종사자 수

2) 10인 이상 사업체, KSIC 중분류 산업 기준, ( ) 안은 KSIC 소분류 산업 기준

제조업종의 자본생산성이 높다는 것은 투입된 자본 대비 부가가치 창출 효율이 높다는 의미이다. 석유정제 11.8%, 자동차 8.8%, 의약품 4.6%로 자본생산성이 증가하였으나, 제조업 전체적으로는 0.7% 감소하였다.

〈표 Ⅲ-7〉 주요 제조업종별 자본생산성 및 증감률

단위: 단위비율, %

2019			2023			연평균 증감률		
1	담배	2.1	1	의복·모피	2.0	1	석유정제	11.8
2	의복·모피	2.0	2	자동차	1.3	2	자동차	8.8
3	가죽·신발	1.7	3	석유정제	1.3	3	의약품	4.6
				~ · ~ · ~				
	제조업	0.9		제조업	0.9		제조업	-0.7

자료: 국가데이터처, 광업제조업조사

주: 1) 자본생산성=부가가치/유형자산 연말잔액

2) 10인 이상 사업체, KSIC 중분류 산업 기준, ( ) 안은 KSIC 소분류 산업 기준

경영성과 분석에서 출하액 기준으로 제조업 주력 업종의 외형적 성장이 지속되고 있다. 제조업 전체로는 연간 6.6% 증가한 가운데 전기장비 13.7%, 의약품 11.2%, 자동차 10.2%, 등이 높은 성장세를 실현하였다.

〈표 Ⅲ-8〉 주요 제조업종별 출하액 및 증감률

단위: 10억 원, %

2019			2023			연평균 증감률		
1	전자제품	254,049	1	자동차	289,075	1	전기장비	13.7
2	자동차	195,774	2	전자제품	269,301	2	의약품	11.2
3	화학제품	153,757	3	화학제품	188,532	3	자동차	10.2
	~ · ~ · ~							
	제조업	1,541,790		제조업	1,988,894		제조업	6.6

자료: 국가데이터처, 광업제조업조사

주: 10인 이상 사업체, KSIC 중분류 산업 기준, ( ) 안은 KSIC 소분류 산업 기준

부가가치액은 연간 4.8% 증가한 가운데 석유정제 16.0%, 의약품 14.6%, 전기장비 12.5% 등이 높은 성장세를 실현하였다.

〈표 Ⅲ-9〉 주요 제조업종별 부가가치액 및 증감률

단위: 10억 원, %

2019			2023			연평균 증감률		
1	전자제품	136,796	1	전자제품	114,161	1	석유정제	16.0
2	(반도체)	87,225	2	자동차	81,485	2	의약품	14.6
3	자동차	55,393	3	(반도체)	72,429	3	전기장비	12.5
	~ · ~ · ~							
	제조업	557,015		제조업	672,555		제조업	4.8

자료: 국가데이터처, 광업제조업조사

주: 10인 이상 사업체, KSIC 중분류 산업 기준, ( ) 안은 KSIC 소분류 산업 기준



04 해외 주요국들은 어땠을까?

1. 해외 주요국들의 산업정책과 정부의 역할은?

주요국의 산업정책 현황을 보면, 경제의 복원력 확보와 장기적인 성장 기반 마련을 위해 제조업의 중요성을 재조명하고, 코로나19 이후 지정학적 리스크 확대와 글로벌 패권 경쟁·블록화에 대응하여 산업정책을 강화하고 있다.

향후 해외 주요국에서 정책 수단을 다양화하고 자국 중심주의를 강화함에 따라 글로벌 산업 지형 변화가 가속화될 것으로 예상된다. 글로벌 산업정책은 국가의 소득 수준과 경제적 목표에 따라 다양한 방식으로 추진되고 있으며, 첨단산업 육성, 공급망 안정화, 디지털·녹색 전환, 경제 안보 등을 추구하는 입법 활동이 확대되고 있음을 알 수 있다.

〈표 Ⅳ-1〉 주요 산업·기술 정책 현황

공급망 안정화	• (EU) 핵심원자재법(CRMA) • (美) 청정구매법(BCA), 미국산우선구매법(BAA) • (韓) 경제 안보를 위한 공급망 안정화 지원 기본법, 소재부품장비산업 경쟁력 강화 및 공급망 안정화를 위한 특별법
디지털·녹색 전환	• (美) 인플레이션감축법(IRA) • (EU) Twin Transition(DX·GX), 탄소중립 산업법(NZIA), 디지털 시장법(DMA), ManufacturingX • (日) 탈탄소성장형 경제구조로의 원활한 이행의 추진에 관한 법률(GX) • (韓) 기후위기 대응을 위한 탄소중립·녹색성장 기본법
경제 안보·무역규제	• (EU) 탄소국경조정제도(CBAM) • (美) 청정경쟁법(CCA) • (일) 경제 안전보장을 위한 추진법 • (한) 산업기술 유출 방지 및 보호를 위한 특별법

2. 최근 5년간 주요국들의 변화는?

최근 5년간 주요국 산업구조의 변화를 살펴보면 코로나19 이전보다 선진권 국가들의 제조업 비중은 전반적으로 감소 경향을 보였고, 신흥권 국가들의 제조업 비중은 증가하는 추세이다. 선진국 평균은 15.8%이고, 일본과 독일은 20% 남짓으로 신흥권 국가들이 제조업 강국임을 알 수 있다.

주요국의 산업정책이 강화되고 있지만 신흥국인 중국(0.7%p), 대만(3.6%p)의 제조업 비중은 증가하였다. 반면에 미국, 프랑스 등은 10% 내외에서 작은 비중을 차지하고 있다.

〈표 Ⅳ-2〉 주요국의 제조업 비중 변화 (2019-2022)

	제조업 비중(%)				제조업 비중 변화(%p)			
	2019	2020	2021	2022	'19-'20	'20-'21	'21-'22	'19-'22
세계 평균	12.41	12.43	12.57	12.50	0.02	0.14	-0.07	0.09
선진국 평균	15.48	15.24	15.88	15.76	-0.24	0.64	-0.12	0.28
미국	11.01	10.77	10.76	10.70	-0.24	-0.01	-0.06	-0.31
일본	21.09	20.71	21.81	21.87	-0.38	1.09	0.06	0.78
독일	20.40	19.53	20.71	20.29	-0.86	1.18	-0.43	-0.11
신흥국 평균	21.80	22.58	23.47	23.25	0.78	0.89	-0.22	1.45
중국	27.90	28.00	28.49	28.59	0.10	0.49	0.10	0.70
대만	31.91	33.18	35.81	35.52	1.27	2.63	-0.29	3.61

자료 : UNIDO DB(2025.6) 이용하여 산업연구원 작성
주 : 세계 평균은 UNIDO(세계 공업 기구) database에 포함된 국가들의 제조업 비중 평균을 의미

05 앞으로 우리는 어떻게 대응해야 할까?

세계적으로 미국·EU·중국 등 주요국은 첨단산업 육성, 공급망 재편, 디지털·녹색 전환을 앞세워 자국 중심의 산업정책을 강화하고 있다. 향후 주요국의 보호무역 강화와 기술 패권 경쟁이 가속화되면서 글로벌 산업 지형은 더욱 요동칠 전망이다. 이에 따라 한국은 세 가지 방향에서 대응 전략을 마련해야 한다.

첫째, 자동차, 조선, 석유화학, 철강, 반도체 등 5대 주력산업의 성장 둔화를 극복하기 위해 첨단·고위기술 제품군으로 생산구조를 고도화하여 주력산업의 경쟁력을 강화해야 한다. 이러한 전략은 기존 주력산업의 체질을 개선하고, 글로벌 공급망 불안 속에서도 안정적인 경쟁 우위를 확보하는 데 기여할 수 있다.

둘째, 질 좋은 일자리를 창출하기 위해 인적자본 및 자동화·AI화자본 투자를 확대하여 고부가가치 산업 생태계를 구축해야 한다. 동시에 산·학·연·관 협력 기반을 강화하고, 기술·자본·인력의 혁신 역량을 제고해야 한다.

이러한 전략은 고용의 질을 높이고, 미래 인재 육성과 기업 경쟁력 제고라는 두 가지 성과를 동시에 달성할 수 있다.

셋째, 식료품, 의약품, 의료·정밀·광학 등 최근 고성장세를 보이는 산업의 수출 확대, 규제 혁신, R&D 및 인프라 투자를 통해 새로운 성장축으로 육성해야 한다. 아울러 로봇·차세대 모빌리티·바이오헬스·수소 등 기술 융합형 신산업에 선제적으로 투자해 新성장 모델을 확보해야 한다. 이러한 전략은 산업 다각화와 혁신산업 발굴을 기반으로 한국 경제의 장기적 성장 동력을 강화하는 데 이바지할 것이다.

결국 주력산업의 경쟁력 강화, 혁신 생태계 구축, 신성장 산업 육성이라는 다층적 전략이 유기적으로 결합될 때, 한국 제조업은 고도화된 산업구조와 글로벌 경쟁력을 동시에 달성할 수 있다.

생성형 AI 시대, 데이터 윤리의 새로운 기준

김윤명 | 디지털정책연구소 소장



데이터 윤리는 데이터 처리 과정에서 발생할 수 있는 윤리적 리스크를 최소화하고, 문제가 발생할 경우 이를 적절히 해결할 수 있는 시스템을 갖추는 일이다. 완전한 차단을 목표로 하기보다는 사전 예방 - 신속 대응 - 재발 방지의 선순환을 설계하는 유연성이 핵심이다. 특히 생성형 AI의 확산으로 데이터의 수집·가공·활용이 고도화된 오늘, 데이터 윤리는 규정 준수를 넘어 신뢰를 만드는 조직 역량이 되었다.

| 1 | 왜, 데이터 윤리인가?

데이터 윤리는 데이터가 개인, 기업, 사회에 미치는 영향을 고려하여 법적, 윤리적 기준을 준수하며 데이터를 다루는 방식이다. 데이터가 인간의 프라이버시와 권리를 침해할 수 있는 잠재적인 위험을 내포하고 있기 때문에, 데이터 윤리의 중요성은 날로 커지고 있다. 예를 들어, 기업이나 정부가 데이터를 수집하고 분석할 때, 이를 어떻게 활용할지에 대한 윤리적 고민이 부족하면 사회적 신뢰를 잃을 수 있다.

이러한 위험은 누적되어 법적·평판 리스크로 전이되기 쉬우므로, 데이터 윤리는 일회성 준수 활동이 아니라 지속적 관리와 학습의 대상이다. 데이터 리터러시 역시 단순한 도구 사용법을 넘어, 데이터를 책임 있게 읽고 쓰는 태도와 규범을 포함한다. 생성형 AI 환경에서는 이 윤리적 이해가 모델 설계·데이터 선택·출력 점검 전 과정의 기본 전제다.

| 2 | 데이터 생애주기와 윤리

데이터 윤리는 데이터의 생애주기 전반에 걸쳐 발생할 수 있는 윤리적 문제를 관리하는 과정이다. 데이터의 수집, 저장, 가공, 분석, 활용, 폐기 단계마다 다양한 윤리적 이슈가 발생할 수 있기 때문에, 이를 관리하는 체계적 접근이 필수적이다. 예를 들어, 데이터 수집 단계에서는 이용자의 동의를 충분히 받고, 수집하는 데이터의 종류와 목적을 명확하게 밝혀야 한다.

데이터 가공 단계에서는 편향된 데이터가 포함되지 않도록 주의하고, 데이터 분석 단계에서는 공정하고 객관적인 방법으로 분석을 수행해야 한다. 데이터 윤리는 이러한 과정을 통해 발생할 수 있는 윤리적 리스크를 최소화하고, 문제가 발생할 경우 이를 적절히 해결할 수 있는 시스템을 갖추는 것이다. 또한, 완전한 차단을 목표로



하기보다는 문제가 발생하지 않도록 예방하고, 발생 시 신속하게 대응할 수 있는 유연성을 지니는 것이 중요하다.

| 3 | 데이터의 윤리적 이슈

(1) 개인정보 보호

개인정보 보호는 데이터 윤리에서 가장 중요한 문제 중 하나다. 개인의 프라이버시를 침해하거나, 동의 없이 민감한 데이터를 수집하는 것은 심각한 윤리적 문제를 야기한다. GDPR(General Data Protection Regulation, 일반 데이터 보호 규정)과 같은 법적 규제는 개인정보 보호를 강화하고 있지만, 각 기업이 법적 요구사항을 넘어 윤리적 책임을 다하는 것이 중요하다. 예를 들어, 기업이 사용자 데이터를 수집할 때, 그 데이터가 어떻게 활용될지에 대해 명확히 고지하고 동의를 얻어야 하며, 수집된 데이터를 안전하게 보호해야 한다. 또한, 불필요하게 데이터를 보유하거나 과도한 데이터를 수집하지 않도록 주의해야 한다. 데이터 유출이나 오용이 발생하면, 이는 법적 제재와 함께 기업의 신뢰도와 평판에 치명적인 영향을 미친다.

(2) 데이터의 편향성

데이터의 편향성 문제는 데이터 윤리에서 매우 중요한 이슈다. 데이터가 특정 집단이나 특성을 과도하게 반영



하거나, 특정 집단을 배제하는 경우, 이로 인해 도출된 분석이나 예측은 왜곡될 수 있다. 예를 들어, 인공지능이나 머신러닝 모델이 학습하는 데이터가 특정 성별, 인종, 지역 등의 특성에 치우쳐 있다면, 해당 모델은 편향된 결정을 내리게 될 것이다. 이는 차별적 결과를 초래할 수 있으며, 사회적 불평등을 심화시킬 수 있다. 따라서 데이터 수집 및 분석 과정에서 데이터의 균형과 대표성을 고려해야 한다. 이를 위해서는 다양한 집단의 데이터를 고르게 수집하고, 분석 과정에서 공정성과 객관성을 확보하는 노력이 필요하다.

(3) 알고리즘의 투명성

알고리즘의 투명성은 데이터 윤리에서 또 다른 핵심 요소다. 데이터 분석이나 인공지능 알고리즘을 사용할 때, 그 과정이 투명하게 공개되지 않으면 결과에 대한 신뢰를 얻기 어렵다. 특히, 중요한 사회적 결정이나 의사 결정을 자동화하는 시스템에서는 알고리즘의 작동 원리와 그 데이터 처리 과정을 명확히 설명할 수 있어야 한다. 예를 들어, 의료 분야에서 AI 시스템이 환자 치료 방침을 결정한다면, 해당 시스템이 어떻게 데이터를 분석하고 어떤 기준으로 결정을 내리는지에 대한 투명성이 필요하다. 이를 통해 사용자는 시스템의 결정을 이해하고, 필요할 경우 이를 수정하거나 재조정할 기회를 가질 수 있다. 알고리즘의 불투명성은 불신을 초래하고, 잘못된 결정을 내릴 위험을 높인다.

(4) 데이터 보안

데이터 보안은 데이터 윤리의 또 다른 중요한 요소로, 데이터가 안전하게 보호되지 않으면 해킹, 유출, 변조 등의 위험에 노출될 수 있다. 개인정보 보호와 관련된 법적 요구사항은 물론, 기업이나 기관이 수집한 모든 데이터는 적절히 암호화되고, 접근 제어가 되어야 한다. 데이터 유출이 발생하면, 기업은 법적 책임을 지게 될 뿐만 아니라 고객의 신뢰를 잃을 수 있다. 또한, 데이터의 보안이 취약하면, 해커가 민감한 정보를 악용하여 사회적, 경제적 피해를 발생시킬 수 있다. 따라서 기업은 데이터 보안을 강화하기 위한 기술적, 관리적 조치를 마련하고, 이를 지속적으로 점검해야 한다. 이는 단순히 법적 준수를 넘어서, 기업의 사회적 책임을 다하는 과정이다.

(5) 데이터 공유와 접근성

데이터 윤리에서 또 다른 중요한 이슈는 데이터의 공유와 접근성이다. 데이터를 공유하는 것은 연구와 개발의 효율성을 높이고, 사회적 가치를 창출할 기회를 제공한다. 그러나 데이터를 공유할 때는 개인정보 보호와 같은 법적, 윤리적 요구사항을 충족시켜야 한다. 예를 들어, 공공기관이 데이터를 공개할 때, 그 데이터에 민감한 정보가 포함되지 않도록 주의해야 하며, 공개된 데이터가 불법적인 목적으로 악용되지 않도록 관리해야 한다.
데이터 접근성 또한 중요한 윤리적 고려사항으로, 데이터를 누구나 쉽게 접근할 수 있도록 하되, 이를 제어할 수 있는 적절한 관리 체계를 마련해야 한다. 데이터의 자유로운 접근이 민주화를 촉진할 수 있지만, 이를 잘못 관리하면 개인정보 유출이나 데이터 왜곡 등의 문제가 발생할 수 있다.

| 4 | 생성형 AI와 데이터 윤리의 관계

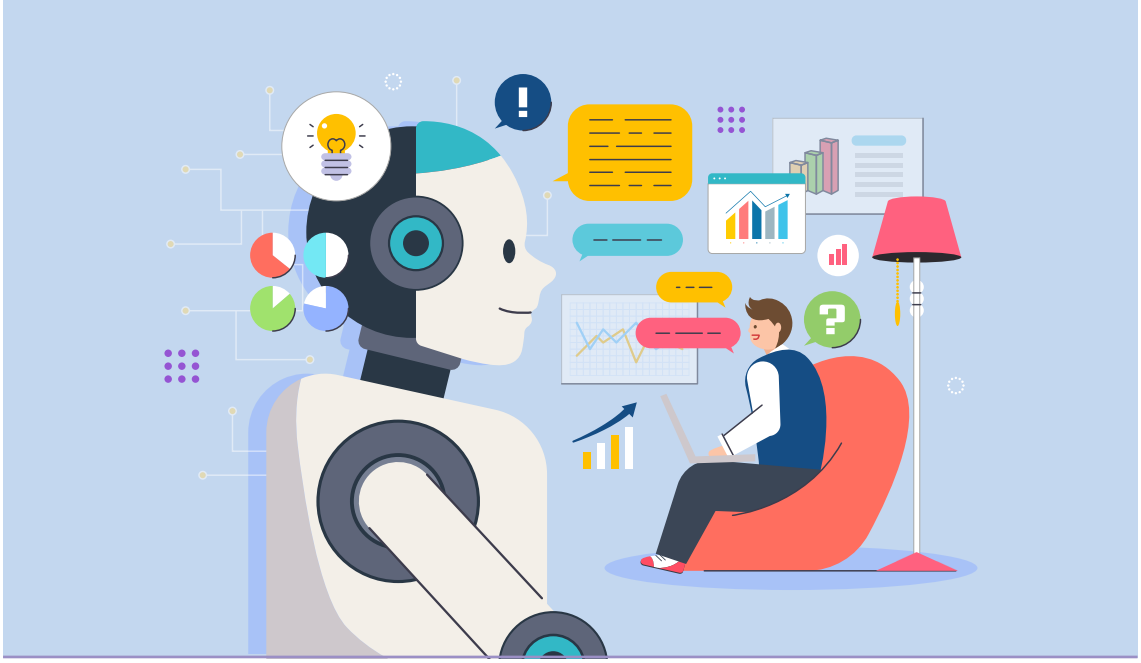
생성형 AI는 방대한 웹·공공·민간 데이터를 섞어 배우고 답을 만들어내기 때문에, 데이터 윤리는 선택이 아니라 처음부터 설계에 들어가야 하는 기본 원칙이 되었다. 무엇보다 데이터가 어디서 왔는지, 어떤 권리·동의를 요구하는지 분명히 하고 모델이 학습 과정에서 기억한 민감정보가 의도치 않게 다시 드러나지 않도록 통제해야 한다.

또한, AI는 그럴듯하지만 틀린 정보를 제공할 수 있으므로 근거 표기와 신뢰도 표시를 기본값으로 삼고, 학습 데이터의 편향이 결과의 차별로 이어지지 않도록 공정성 지표와 영향평가를 운영하는 것도 고려해야 할 것이다. 데이터 소싱과정에서 계약도 중요한 요소이다. 특히, 데이터 계약을 체결할 때 실무적으로 고려해야 할 주요 사항들은 데이터의 법적, 기술적, 운영적 측면을 모두 아울러야 한다는 점이다. 이를 통해 계약 당사자 간의 권리와 의무를 명확히 하고, 불필요한 리스크를 최소화할 수 있다.

| 5 | 맺음말: 시대의 흐름 속에서 국가데이터처의 역할

최근 통계청의 확대·개편으로 국가데이터처가 출범하며, 국가 데이터 거버넌스의 중심축이 재정렬되고 있다. 이는 공신력 있는 통계를 넘어, 표준·품질·보안·윤리를 통합하는 데이터 컨트롤타워의 등장을 뜻한다. 이러한 변화 속에서 국가데이터처는 ① 생성형 AI 친화적 메타데이터와 지표 설명서의 표준화, ② 안전한 분석을 위한 안심구역·샌드박스의 고도화, ③ 기관별 윤리·공정성 성숙도 진단과 사례집 확산을 주도하는 촉진자 역할을 맡을 수 있다.

AI 3대 강국을 위한 데이터 정책은 '모두의 AI'나 'AI 기본사회' 구현을 위한 기본적인 요소이기도 하다. 시시각각 변화하는 AI 발전 속도와 신뢰를 동시에 요구하는 시대, 데이터 윤리는 혁신의 제동이 아니라 방향과 충격을 잡아주는 규범이다. 이 원리가 현장의 습관과 제도로 자리 잡을 때, 생성형 AI는 더 안전하고 공정하게 국민의 삶을 개선할 것이다.



인공지능 시대, 엔터테인먼트는 어떻게 진화하는가?

황서이 | 중부대학교 미디어커뮤니케이션학과 교수



1
들어가며

인공지능(AI, Artificial Intelligence)은 더 이상 기술의 영역에 머물지 않는다. 오늘날 AI는 콘텐츠의 기획과 제작을 넘어 유통과 소비의 전 과정에 깊숙이 스며들며, 엔터테인먼트 산업의 문화적 인프라이자 새로운 창작 언어로 자리 잡았다. 음악, 드라마, 영화, 광고, 마케팅 등 우리가 즐기는 거의 모든 무대 뒤에는 이제 AI가 존재한다.

SNS에서는 '김햄찌', '냠냠이' 같은 AI 캐릭터들이 수만 명의 팔로워를 모으며, 단순한 오락을 넘어 정서적 교감과 위로의 주체로 기능하고 있다. 신세계 까사(SHINSEGAE CASA)는 이러한 흐름을 반영해 AI 캐릭터를 브랜드 캠페인에 도입함으로써 초상권 부담을 줄이고, 감성적 소통을 강화하는 새로운 방식을 제시하였다.

방송과 광고, 영화 산업에서도 AI의 활용은 빠르게 확산하고 있다. MBC 〈놀면 뭐하니〉의 '서울가요제' 인트로 영상은 1980년대와 현재를 잇는 콘셉트를 구현하기 위해 AI로 제작하였다. 윤여정 배우의 젊은 시절을 복원한 디에이징(De-aging) 기술, 고(故) 임윤택(울랄라 세션)과 터틀맨(거북이), 김현식을 AI로 재현한 공연, 지드래곤이 AI를 활용해 제작한 뮤직 비디오 〈Home Sweet Home〉을 우주로 송출한 사례 등은 기술이 예술의 시간과 감정을 새롭게 복원하고 확장하는 과정을 보여준다.

AI는 인간의 예술을 대체하기보다 감성을 확장하고 기억을 재구성하며, 예술의 지속성을 이어주는 매개체로 진화하고 있다. 기술의 영역에서 출발한 AI가 이제는 창작과 감동의 언어로 자리 잡으며, 엔터테인먼트 산업의 패러다임을 근본적으로 바꾸고 있다.

2
AI가 재구성하는 기획·제작·유통·소비

AI는 이제 콘텐츠 산업의 전 과정을 새롭게 쓰고 있다. 대규모 언어 모델(LLM, Large Language Model)은 언어와 문화의 장벽을 낮추며 번역과 자막을 자동으로 제안하고, 확산모델(Diffusion Model)은 텍스트 입력만으로 고품질 이미지를 생성해 제작 효율을 높인다. 여기에 보이스 클로닝(Voice Cloning)과 모션 리타깃팅(Motion Retargeting) 기술이 결합하면서 배우와 가수의 목소리, 표정, 동작까지 정교하게 재현된다. 그 결과, 글로벌 보편성과 지역 특수성을 동시에 구현하는 맞춤형 창작의 시대가 열리고 있다.





국내에서도 이러한 변화는 이미 현실이 되었다. KB라이프생명은 디에이징 기술을 활용해 젊은 시절의 윤여정 배우를 복원하며 세대를 잇는 감성 광고를 선보였다. 디즈니플러스 〈카지노〉에서는 배우 최민식이 AI 보이스 디에이징 기술을 적용해 30대 시절의 얼굴과 목소리를 되살렸다. 이 기술은 단순히 외형을 젊게 만드는 수준을 넘어, 배우의 실제 음성을 기반으로 과거 작품의 발성과 감정선을 학습해 자연스러운 젊은 목소리를 복원하는 방식이다. 제작진은 여러 버전의 음성을 실험한 끝에 연기 톤과 감정의 흐름이 가장 자연스러운 모델을 선택해 최종 장면에 반영하였다. JTBC 〈웰컴투 삼달리〉에서는 고(故) 송해의 목소리와 표정을 딥페이크(Deepfake) 기술로 복원해 "전국~ 노래자랑!"의 상징적 멘트를 다시 무대 위로 불러냈다. 이 사례는 기술이 인간의 예술을 단순히 복제하는 것이 아니라, 시간과 감정을 복원하며 예술의 지속성을 확장하는 과정임을 보여준다.

AI의 영향력은 이제 엔터테인먼트 산업 전반으로 확산하고 있다. 하이브(HYBE)는 AI 오디오

기업 수퍼톤(Supertone)의 기술을 활용해 아티스트의 음성을 다국어로 구현하고, 신곡을 전 세계에 실시간 발매함으로써 글로벌 팬과의 소통 방식을 혁신하였다. SM엔터테인먼트의 버추얼 아티스트 나이비스(Naevis)는 AI 보이스와 시각특수효과(VFX)를 결합해 현실과 가상의 경계를 허무는 공연을 선보였다.

AI는 이제 기획·제작뿐 아니라 유통과 소비까지, 콘텐츠의 전 생애주기를 재편하고 있다. 기획 단계에서는 시장 데이터를 분석해 흥행 가능성을 예측하고, 제작 단계에서는 영상·음향 합성을 통해 제작 시간을 절감하며, 유통 단계에서는 개인 맞춤형 추천으로 소비자 경험을 정교하게 확장한다. 이러한 기술의 도입은 단순한 효율화를 넘어 제작비 절감, 다언어 동시 유통, 글로벌 소비 확대 등 산업 구조의 근본적 변화를 이끌고 있다.

무엇보다 AI는 창작의 문턱을 낮추었다. 전문 장비나 대규모 인력이 없어도 텍스트 한 줄의 프롬프트만으로 시나리오가 완성되고, 음악이 작곡되며, 영상이 제작된다. 이제 팬은 완성된 콘텐츠의 수동적 소비자가 아니라, 리믹스 영상이나 팬아트, 버추얼 무대를 직접 제작하는 참여형 창작자로 진화하고 있다.

3 엔터테크가 이끄는 새로운 무대

AI는 이제 단순한 제작 도구가 아니다. 콘텐츠의 기획부터 제작, 유통, 소비까지 전 과정을 변화시키는 강력한 동력으로 자리 잡았다. 그러나 생성형 AI 콘텐츠 시대의 경쟁력은 단순히 기술을 보유했는가에 달려 있지 않다. 오히려 기술을 얼마나 깊이 이해하고, 그것을 창작의 언어로 재해석하느냐가 중요하다. 결국, 수용자의 몰입 경험 속에서 AI를 자연스럽게 녹여낼 수 있는 능력, 그것이 진정한 경쟁력이다.

이러한 변화를 이끄는 핵심 개념이 바로 엔터테크(EnterTech)다. 엔터테인먼트(Entertainment)와 테크놀로지(Technology)의 합성어인 엔터테크는 음악, 영상, 공연 등 다양한 콘텐츠에 AI, 가상현실(VR), 증강현실(AR), 메타버스(Metaverse), 버추얼 휴먼(Virtual Human) 등 첨단 기술을 결합해 새로운 부가가치를 창출하는 산업을 뜻한다. 즉, 엔터테크는 단순히 효율을 높이는 기술이 아니라, 사람과 사람을 잇는 '관계의 기술', 그리고 감정의 깊이를 더하는 '경험의 기술'이다. OTT 서비스 경쟁이 치열해지고, 팬덤 경제가 확산되며, 창작자와 소비자의 상호작용이 커질수록 엔터테크의 중요성은 더욱 두드러지고 있다.

무엇보다 그 변화는 이미 일상에서 현실이 되고 있다. 지드래곤은 AI로 제작한 뮤직비디오를 우주로 송출하며 창작의 경계를 확장했고, JTBC 〈얼라이브(ALIVE)〉는 세상을 떠난 가수의 목소리를 다시 무대 위로 불러냈다. 또한, K-pop 아티스트들은 AI 오디오와 버추얼 캐릭터를 활용해 수용자와의 소통 방식을 다층적으로 확장하고 있다. 이처럼 엔터테크는 효율을 높이는 기술이 아니라 감성을 확장하는 기술이다. 기술이 예술의 감정을 다시 불러내는 순간, 우리는 새로운 무대의 탄생을 목격하게 된다.

4
나오며

2022년 말, ChatGPT의 등장은 전 세계에 생성형 AI의 가능성과 충격을 동시에 안겼다. 불과 2년 만에 텍스트, 이미지, 음악, 영상 등 다양한 형식의 콘텐츠가 AI를 통해 제작되며, 인간의 창작물과 거의 구분되지 않을 만큼 정교해졌다. 이제 생성형 AI는 보조 도구를 넘어 콘텐츠 제작의 실질적 주체로 자리 잡았다.

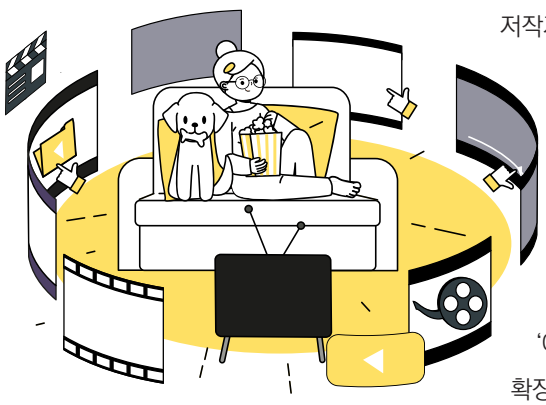
전 세계적으로 음악, 영화, 드라마, 방송·영상, 광고, 마케팅 등 엔터테인먼트 전반에서 AI는 기획부터 제작, 유통, 소비, 그리고 팬과의 소통에 이르기까지 전 과정을 아우르고 있다. 그러나 기술 발전이 반드시 긍정적인 결과만을 가져온 것은 아니다.

2023년, 미국작가조합(WGA)과 배우조합(SAG-AFTRA)은 연대 파업에 돌입하였다. 표면적으로는 처우 개선과 OTT 수익 배분이 주요 쟁점이었지만, 그 근본에는 AI의 도입이 창작자와 배우의 일자리를 위협할 수 있다는 불안이 자리하였다. 장기간의 제작 중단과 막대한 손실로 이어진 이 사건은 AI가 엔터테인먼트 산업의 노동 구조와 창작 환경에 미치는 파급력을 단적으로 보여준 사례다.

같은 해 4월, 드레이크(Drake)와 위켄드(The Weeknd)의 곡 〈Heart on My Sleeve〉가 AI가 생성한 가짜 음원으로 판명되며 주요 플랫폼에서 삭제된 사건도 있었다. 이 일은 플랫폼, 창작자, 기술 개발자 간의 법적·윤리적 책임 문제를 본격적으로 제기하였다. 이후 유니버설 뮤직(UMG)과 워너뮤직(Warner Music) 등 글로벌 음반사는 무단 사용을 차단하기 위해 핑거프린팅(Fingerprinting) 기술을 도입하며, 기술과 협력이 결합한 새로운 대응 체계를 마련하고 있다.

이러한 사례는 한 가지 사실을 명확히 보여준다. 기술의 속도가 사회적 이해의 속도를 앞지를 때, 산업은 방향을 잃는다. 이는 기술이 너무 빠르게 발전할 때, 이를 이해하고 규율할 법·윤리·교육 같은 사회적 기반이 따라가지 못하면 산업 전체가 '어디로 가야 할지 모르는 상태'에 빠진다는 뜻이다.

지속 가능한 AI 콘텐츠 생태계를 위해서는 기술 발전만큼 윤리적 감수성과 사회적 기반, 법적 제도가 함께 성장해야 한다. 창작자에게는 AI의 작동 원리와 영향력을 이해하는 윤리 교육이, 수용자에게는 AI 리터러시가 필요하다. '이 콘텐츠는 AI로 제작되었습니다.'와 같은 명시를 통해 투명성을 확보하고, 뮤직비디오나 버추얼 인터뷰 등에서는 AI 개입 여부를 명확히 공개해야 한다. 또한 'AI는 도구인가, 창작자인가?', 'AI가 학습한 데이터의 원저작자에게 어떤 보상이 주어져야 하는가?'와 같은 쟁점에 대한 사회적 합의도 시급하다.



최근 GPT-4o 등 생성형 AI는 개인 이미지를 캐릭터화하는 등 창작의 민주화를 이끌고 있지만, 동시에 저작권·초상권·책임 문제를 새롭게 제기하고 있다. AI는 이제 전문가의 도구가 아닌 일상의 언어가 되었으며, 그만큼 사회적 이해와 제도적 대응이 더욱 정교해져야 한다. 결국, AI와 엔터테인먼트의 미래는 기술의 속도보다 '이해와 책임의 속도'에 달려 있다. 기술이 인간의 상상력과 윤리를 함께 확장할 때, 진정한 '엔터테크 시대'가 완성될 것이다.

통계로 바라보는 세상이야기

신동헌 | 도서출판 지일북스 대표

광복 80년, 독립유공자들의 오늘은?

국가보훈부의 2024년 「국가보훈대상자 생활실태조사」에 따르면, 독립유공자 가구의 가구원 수는 48.6%로 2인 가구가 가장 많았으며, 1인 가구가 32.9%로 그 뒤를 이었습니다. 이어 3인 가구가 10.8%, 4인 가구가 4.4%, 5인 이상 가구가 3.3%를 차지했습니다. 독립유공자 가구는 대체로 1~2인 가구가 절반 이상으로 나타났습니다. 특히 이들은 몸이 아파 집안일을 부탁해야 할 경우, 집안일 도움을 요청할 수 있는 사람이 있다고 응답한 비율이 84.1%로 가장 높았고, 이야기 상대가 필요한 경우, 도움을 요청할 수 있는 사람이 있다고 응답한 비율은 80.4%로 뒤를 이었습니다. 반면, 갑자기 돈을 빌려야 할 경우, 도움을 요청할 수 있는 사람이 있다고 답한 비율은 57.3%로 상대적으로 낮은 편이었습니다.

독립유공자, 지원정책 긍정평가 37.2%

독립유공자분들의 마음속 자긍심은 어느 정도일까요? 독립유공자의 자긍심 정도를 살펴본 결과, 전체의 68.0%가 '매우 그렇다(17.7%)' 또는 '그렇다(50.3%)'고 답해 독립유공자 대다수가 강한 자긍심을 가지고 있었습니다. '그렇지 않다'와 '전혀 그렇지 않다'는 합쳐서 5% 미만으로 매우 적었습니다. 교육, 취업, 의료, 생활 안정 등 다양한 지원정책에 대해 얼마나 만족하고 계실까요? '매우 만족'은 1.4%, '만족'은 32.7%로, 전체의 34.1%가 긍정적으로 평가했습니다. 이는 현재의 지원 정책에 대해 전반적으로 만족하며, 그 혜택이 실제 삶에 의미 있게 작용하고 있음을 보여줍니다. 광복 80년, 오늘 우리가 누리는 자유와 번영은 독립유공자분들의 희생과 헌신 위에 세워졌음을 기억해야 합니다.

우리 청년들이 바라는 미래는?

국무조정실의 「2024년 청년삶실태조사」에 따르면, 청년들이 현재 상황에서 바라는 삶의 요소 중 가장 중요하다고 생각하는 요소는 '내가 원하는 일자리(95.9%)'였습니다. 이어 '좋은 사람들과의 관계(94.7%)', '높은 소득과 많은 자산(93.0%)'이 그 뒤를 이었으며, '연애(78.3%)'와 '결혼(74.4%)'도 중요한 요소로 꼽혔습니다. 실현 가능한 지에 대해서는, 응답자의 대다수가 '어느 정도는 실현할 수 있지만, 완벽하게 실현할 수 없다'고 답했으며, 그 비율은 2022년 87.6%에서 2024년 86.7%로 비슷한 수준을 보였습니다. 한편, '전혀 실현할 수 없다'는 응답은 2022년 5.2%에서 2024년 7.6%로 증가한 반면, '완벽하게 실현할 수 있다'는 응답은 2022년 7.2%에서 2024년 5.7%로 감소했습니다. 일정 부분의 성취 가능성은 인정하지만, 원하는 삶을 실현하기에는 여전히 여러 제약이 따른다고 인식하는 것을 보여줍니다.

미래를 위해 가장 중요한 것은?

국무조정실 「2024년 청년삶실태조사」에 따르면, 청년들이 바라는 미래를 실현하기 위해 가장 중요한 요소로 '나의 노력(97.8%)'과 '좋은 사람들과의 관계(94.8%)'가 가장 많이 꼽혔습니다. 이어서 '정부 정책 지원(89.3%)', '나의 교육 수준(87.4%)', '부모의 경제력(83.0%)'이 중요하다고 응답했으며, 상대적으로 '부모의 교육 수준(61.4%)'은 중요도가 낮게 나타났습니다. 청년들은 개인적 노력뿐 아니라 사회적·제도적 지원도 함께 필요하다고 인식하고 있습니다. 청년들이 가장 필요하다고 생각하는 정책은 주거 정책(45.7%) 이었습니다. 이어 '일자리 정책(33.5%)', '복지·문화 정책(14.1%)', '교육 정책(5.5%)', '참여·권리 정책(1.2%)' 순으로 나타났습니다. 청년들은 주거와 일자리 관련 정책을 가장 필요로 하며, 복지·교육·참여 정책은 상대적으로 낮은 우선순위를 두고 있는 것으로 나타났습니다.

청년 취업 준비 “바로 이곳”

한국고용정보원의 「2023년 청년패널조사」에 따르면, 취업을 위해 가장 많이 선택한 노력은 ‘웹 구직정보 습득 및 민간 취업알선 사이트 구직등록’이 17.9%로 가장 많았습니다. ‘회사에 직접 이력서 제출’이 17.3%, ‘사설학원(어학 등) 교육’이 15.9%로 그 뒤를 이었습니다. 취업 준비자들이 가장 많이 이용하는 취업정보 취득 경로는 ‘민간 취업포털사이트와 앱(사람인, 잡코리아 등)’이 24.0%로 가장 높았고, ‘공공 취업포털사이트 및 앱(워크넷 등)’이 14.0%로 그 뒤를 이었습니다. ‘친구나 선후배 등 지인’을 통한 정보 취득이 10.5%, ‘학교 선생님이나 교수를 통한 경우’가 10.0%였으며, 공공취업알선기관을 통한 정보 취득은 7.2%로 나타났습니다. 취업 정보를 얻기 위해 민간 취업포털 외에도 지인 네트워크나 학교, 공공기관 등을 통해 정보를 얻고 있습니다.



청년 취업, 영향력 높은 스펙은?

취업에 가장 큰 영향을 미칠 것으로 생각한 스펙으로는 ‘업무(직무) 관련 자격증’이 5점 만점에 3.81점으로 1위를 차지했고, ‘컴퓨터 관련 자격증(3.48점)’, ‘학점(3.42점)’, ‘토익·토플 등 공인영어성적(3.38점)’, ‘인턴, 아르바이트 등 직무경험(3.27점)’이 그 뒤를 이었습니다. 실제 직무 수행 능력을 보여줄 수 있는 자격증과 경험이 청년 취업 경쟁력의 핵심 요인으로 볼 수 있습니다. 청년들이 취업을 위해 준비한 스펙 중 '업무(직무) 관련 자격증을 준비한 경험이 있다'고 응답한 비율이 33.2%로 가장 높았고, ‘토익·토플 등 공인 영어성적(31.9%)’, ‘학점(29.9%)’, ‘컴퓨터 관련 자격증(26.1%)’ 순으로 나타났습니다. 또한 ‘영어회화 능력’, ‘공사·공단, 7급·9급 공무원, 교사 임용, 고시 등 시험 준비’ 등 다양한 방식으로 취업을 준비하고 있습니다.

취업자, 평일과 주말 차이는?

2024년 국가데이터처의 「생활시간조사」에 따르면, 취업자의 평균 기상 시각은 평일(월~목) 6시 48분, 금요일 6시 47분으로 큰 차이가 없었습니다. 반면, 주말에는 토요일 7시 25분, 일요일 7시 35분으로 다소 늦게 일어나는 경향을 보였습니다. 2019년과 2024년, 취업자들의 평균 근무 시간에는 어떤 차이가 있을까요? 2024년 취업자의 평일 평균 일 한 시간은 6시간 46분으로, 2019년의 6시간 58분보다 12분 줄어든 것으로 나타났습니다. 이는 지난 5년간 취업자들의 일 한 시간이 다소 단축된 경향을 보여주며 일과 삶의 균형에 긍정적인 변화를 시사합니다. 하루는 어떻게 보낼까요? 주간 36시간 이상 근로하는 취업자들은 하루 평균 11시간 13분을 ‘개인 유지’에 사용했습니다. 이어 5시간 43분을 ‘일’에 사용하고, 3시간 6분을 ‘문화·여가활동’에 사용했습니다.

취업자, 삶과 여가 만족도는?

취업자들은 자신의 삶과 여가에 얼마나 만족하고 있을까요? 취업자의 43.8%가 삶에 ‘만족한다’고 응답하여, 불만족 13.0%보다 매우 높게 나타났습니다. 다만 삶의 만족도보다 여가 만족도의 ‘만족한다’가 7.6%p 낮고 ‘불만족한다’가 9.0%p 높아 여가에 대한 만족도가 상대적으로 떨어짐을 보여줬습니다. 하지만 여가 만족도 역시 만족(36.2%)이 불만족(22.0%)보다는 높아, 삶과 여가에 만족하는 취업자가 더 많음을 보여줬습니다. 앞으로 여가 활동이 더욱 다양해져 취업자의 삶과 여가 만족도가 함께 높아지기를 기대합니다. 취업자들은 바쁜 일상 속에서 각자 필요한 시간을 보내며 하루를 채워갑니다. 어떻게 시간을 관리하고 우선순위를 정하느냐에 따라 일상에서 느끼는 만족도와 삶의 질이 달라질 수 있습니다. 조금 느려도, 완벽하지 않아도 괜찮습니다.

2024년 다문화 수용성 인식 변화



여성가족부의 「2024년 국민다문화수용성조사」에 따르면, 2024년 연령대별 다문화 수용성 지수는 20대가 55.4점으로 가장 높게 나타났습니다. 이어서 30대가 54.8점, 40대가 53.5점, 50대가 53.1점으로 연령이 높아질수록 점수가 소폭 낮아지는 경향을 보였습니다. 60세 이상은 51.1점으로 가장 낮은 점수를 보였습니다. 성인들의 이주민 수용태도를 살펴본 결과, 이주민이 나의 친구가 되는 것에 대해 대체로 긍정적인 인식을 가진 것으로 나타났습니다. '매우 불편하다'고 응답한 비율은 2.7%에 그쳤고, '다소 불편하다'는 20.0%로 나타났습니다. 반면 '별로 불편하지 않다'와 '전혀 불편하지 않다'를 포함한 '불편하지 않다'고 응답한 비율은 77.3%로, 성인 대부분이 이주민과 친구가 되는 것에 대해 긍정적인 태도를 보이는 것으로 나타났습니다.

일상 속 외국인 접촉 경험은?

일상 속 외국인과 이주민 접촉 경험을 살펴보면, 전체 응답자의 과반수 이상(55.6%)이 '가끔 본다'고 응답했습니다. '자주 본다'가 27.4%, '매우 자주 본다'는 6.9%로 나타나, 일상에서 외국인과 일정 수준의 접촉을 경험하고 있음을 알 수 있습니다. 반면 '거의 본 적 없다'는 9.3%, '전혀 본 적 없다'는 0.8%에 그쳤습니다. 다문화 교육의 필요성에 대한 인식을 살펴보면, 응답자 중 44.6%는 '대체로 필요하다'고 응답했으며, 6.6%는 '매우 필요하다' 라고 답했습니다. 반면 '전혀 필요하지 않다'는 2.2%에 불과했습니다. 다문화 사회에 대한 이해와 상호 존중을 증진하기 위해 다문화 교육이 중요함을 알 수 있습니다. 우리 사회가 다문화를 이해하고 포용한다는 것은, 서로의 차이를 존중하며 함께 성장하는 공동체를 만들어 가는 과정입니다.

홈쇼핑에서 해외직구까지, 쇼핑 중

2024년 정보통신정책연구원의 「한국미디어패널조사」에 따르면, 전체 응답자의 74.7%가 전자상거래나 온라인 거래 경험이 있다고 응답했습니다. 이는 국민 4명 중 3명은 전자상거래 및 온라인 거래를 경험한 것을 의미합니다. 유형별 이용 형태를 살펴보면 '온라인 쇼핑몰'이 44.5회로 압도적인 비중을 차지했습니다. 그 뒤를 이어 'TV 홈쇼핑 (14.3회)', '개인 간 거래(8.3회)', '해외직구(7.8회)' 순으로 나타났습니다. 전통적으로 주로 이용하던 홈쇼핑 채널과 더불어 개인 간 거래, 해외직구까지 소비자들의 쇼핑방식이 점점 다변화되고 있음을 보여줍니다. TV 홈쇼핑 구매 경험 추이를 살펴보면, 2022년 41.9%였던 TV 홈쇼핑 구매 경험 비율은 2023년 39.5%로 줄었다가, 2024년 43.7%로 다시 반등했습니다. 이는 TV 홈쇼핑이 꾸준히 소비자층을 확보하고 있음을 보여줍니다.

온라인쇼핑 이용 경험 지속 증가

구매 경험이 있다고 응답한 비율은 2022년 85.8%, 2023년 86.6%, 2024년 88.3%로 지속적인 상승세를 보였습니다. 2024년 해외직구 구매 경험을 살펴보면, 전체 응답자의 26.1%가 구매 경험이 있다고 응답했으며, 73.9%는 경험이 없다고 답했습니다. 구체적인 이용 횟수는 '3회 미만'이 37.7%로 가장 많았고, '3~5회 미만'이 26.9%, '5~10회 미만'이 11.8%로 그 뒤를 이었습니다. 해외직구 이용자는 구매 횟수에 따라 다양한 이용 패턴이 존재함을 보여줍니다. 2024년 개인 간 거래 구매 현황을 살펴보면, 전체 응답자의 26.4%가 구매 경험이 있다고 응답했습니다. 이용 횟수는 3회 미만이 27.7%로 가장 많았고 '5~10회 미만'이 22.5%, '3~5회 미만'이 21.3%로 그 뒤를 이었습니다. 반복적으로 거래를 이어가는 활발한 해외직구 소비자층이 존재함을 보여줍니다.

영화 분야 예술인, 전체 4% 수준



다양한 예술 활동 중 영화 분야에서 주로 활동하는 예술인은 전체 예술인의 약 4~5%에 불과합니다. 2018년 4.2%에서 2021년 5.5%로 소폭 증가했으나, 2024년에는 4.0%로 다시 감소했습니다. 즉, 한국 예술인 중 영화 분야에 집중해 활동하는 사람은 상대적으로 적은 편이며, 최근 몇 년 사이 소폭의 등락을 반복하고 있습니다. 2024년 영화 분야에서 활동하는 예술인들의 주요 데뷔 경로를 살펴보면, 전체의 34.4%가 '학교 및 전문 교육기관'을 통해 데뷔했다고 응답했습니다. 또한, '취업이나 프리랜서 활동 등 관련 분야 경험'이 30.7%, '오디션 등'을 통한 데뷔가 23.7%를 차지했습니다. 그리고 '전시회/연주회/공연/방송 참여'가 9.1%, '신춘문예/대회/공모전 참여'가 7.8%로 나타났습니다. 이들은 정규 교육과 현장 경험, 오디션 등을 통해 입문하고 있습니다.

전업 비중은 줄고, 겸업은 증가세

2018년부터 2024년 영화 분야에서 전업으로 활동하는 예술인을 비중을 살펴보면, 2018년 67.6%에서 2021년 64.6%로 소폭 감소했으며, 2024년에는 61.0%로 계속 감소하는 추세를 보입니다. 반대로 겸업으로 활동하는 영화 예술인은 2018년 32.4%에서 2024년 39.0%로 점차 증가했습니다. 전업만으로는 생계를 유지하기 어렵기 때문에 다른 수익 활동을 병행하는 예술인이 점차 늘고 있음을 보여줍니다. 또한, 지난 1년간 외국에서 영화 예술 활동을 한 경험을 살펴보면, 2018년 14.2%, 2021년 14.8%, 2024년 14.3%로 나타났습니다. 최근 6년간 약 14%대가 꾸준히 유지되고 있으며, 이는 해외 활동 경험이 일부 예술인들에게만 제한적으로 이루어지고 있음을 보여줍니다. 앞으로 더 많은 영화 예술인들이 국제 무대에서 경험을 쌓을 수 있기를 바랍니다.

청소년 인권 관련 기관 인지도는?

한국청소년정책연구원이 발표한 「2024년 아동청소년인권실태조사」에 따르면, 아동·청소년의 인권 관련 협약 및 기관의 인지도는 전반적으로 '이름은 들어 봤지만 구체적으로는 잘 모른다'고 응답한 비율이 높게 나타났습니다. 항목별로 살펴보면, '국가인권위원회'를 '모른다'는 응답은 22.2%로 가장 낮았고, '이름만 들어봤다'는 응답이 55.1%로 가장 많았으며, '실제로 알고 있다'는 비율도 22.7%로 항목 중 가장 높았습니다. 반면, '학생인권조례'는 전체의 41.6%가 '모른다' 라고 답했으며, '이름만 들어봤다'는 응답이 46.0%, '이름도 들어봤고, 알고 있다'는 응답은 12.4%에 그쳤습니다. 즉, 아동·청소년들은 인권 관련 협약이나 기관의 이름은 비교적 익숙하지만, 구체적인 내용이나 의미까지 이해하는 경우는 드문 것으로 나타났습니다.

아동청소년 인권교육 경험 하락세

한국청소년정책연구원 「2024년 아동청소년인권실태조사」에 따르면, 인권교육 경험은 학교 급이 높아질수록 낮아지는 경향이 뚜렷하게 나타납니다. '초등학교'는 2022년 63.8%에서 2023년 61.2%, 2024년 60.1%로 해마다 조금씩 줄어드는 것을 알 수 있습니다. '중학교'는 2022년 45.9%에서 2023년 47.1%로 소폭 증가했으나, 2024년에는 다시 44.5%로 감소했습니다. '고등학교'는 최근 3년간 약 42.0%로 거의 변화가 없었습니다. 반면, '인권의 개념과 인권상황에 대한 이해'에 도움이 되었다고 응답한 비율은 91.3%로 가장 높았습니다. '다른 사람의 인권을 존중하는 생각과 태도'에 도움이 되었다고 응답한 비율이 91.1%로 그 뒤를 이었습니다. '인권 침해와 차별을 극복하기 위한 행동 실천'에는 86.5%가 도움이 되었다고 응답했습니다.

AI를 활용하여 텍스트 데이터 수집 및 분석하기

구자룡 | 밸류바인 대표



우리는 일상에서 혹은 비즈니스에서 수많은 데이터를 생산하고 또 소비하고 있다. 대표적으로 자동차 내비게이션은 나의 차량 이동 정보가 나도 모르게 제공되어 다른 사람의 차량 내비게이션에 사용된다. 반대로 다른 사람의 정보를 내가 이용한다. 정형화된 데이터를 AI가 분석해서 제공해 주기 때문에 가능하다. 온라인 쇼핑몰에서 어떤 제품을 구매하고자 할 때 다른 사람들이 작성한 후기를 보고 구매한다. 내가 제공한 후기는 다른 사람들이 보게 될 것이다. 다만 이 경우는 비정형 텍스트 데이터를 우리가 직접 읽고 판단한다.

텍스트 데이터는 그 자체로 보면 빅데이터이지만, 우리가 직접 하나씩 읽을 때는 스몰데이터다. 고객 리뷰, 블로그 포스트, 소셜 미디어 댓글 등이 비즈니스 관점에서 관심을 가질 만한 텍스트 데이터다. 즉, 이런 데이터를 분석함으로써 고객의 의견과 정서를 파악하여 제품 및 서비스 개선 방향을 찾을 수 있다. 우리 브랜드에 대한 언급을 모니터링하여 브랜드의 평판을 관리하고 위기 상황에 신속하게 대응할 수 있다. 또한, 최신 트렌드를 파악하고 경쟁사의 마케팅 전략, 제품 리뷰 등을 분석하여 경쟁사와의 차별화를 모색할 수 있다.

1. 텍스트 데이터 분석의 새로운 도구 : 생성형 AI

우리는 텍스트 데이터를 분석하고 통찰함으로써 비즈니스의 많은 문제를 데이터에 기반하여 해결할 수 있다. 이를 위해서는 텍스트 데이터가 무엇이고 어떻게 수집하고 분석하는지에 대해 이해해야 한다. 정형 데이터에 대한 분석 기법과 도구에 비해 텍스트 데이터 분석을 할 수 있는 분석 기법과 도구는 최근 들어 정립되었다. 통계분석을 기반으로 하는 모델링의 발전 속도에 비하면 한참 늦은 상황이다. 이유는 자연어 처리에 어려움이 많았기 때문이다. 그러나, 생성형 AI가 등장하면서 너무나 쉽게 이 문제가 해결되고 있다. 생성형 AI가 바로 대규모 언어 모델을 기반으로 자연어 처리를 하기 때문이다. 이제 누구나 생성형 AI를 활용하여 텍스트 데이터 분석을 할 수 있는 시대가 되었다. 생성형 AI가 나를 대신하여 파이썬 코딩을 해주기 때문에 누구나 쉽게 텍스트 데이터 분석을 할 수 있다.

텍스트 마이닝(Text Mining)이라고도 하는 텍스트 데이터 분석(Text Data Analysis)은 비정형 텍스트를 정형으로 변환하여 의미 있는 패턴과 새로운 인사이트를 찾아내는 분석 방법이다. 텍스트 데이터 분석을 통해 기업은 제품 및 서비스의 품질을 향상하고, 고객 만족도를 높이며, 시장에서 경쟁 우위를 확보할 수 있다. 텍스트 데이터와 그 분석의 중요성이 더 커지고 있다.

텍스트 데이터는 비정형 데이터로 그 안에 무엇이 포함되어 있는지 명확하지 않기 때문에 이를 통해 의미 있는 정보를 추출하기 위해서는 체계적이고 효율적인 분석이 필요하다. 대표적인 텍스트 데이터 분석 기법으로 형태소 분석, 감정 분석, 주제 모델링, 텍스트 분류, 클러스터링, 워드 클라우드 시각화 등이 있다. 이를 통해 텍스트 데이터를 정형 데이터로 변환해 통계적 분석이나 머신러닝 모델링에 활용할 수 있으며, 분석 기법을 활용해 유의미한 패턴을 찾을 수 있다.

(1) 형태소 분석

형태소 분석은 텍스트 데이터를 가장 기본적인 의미 단위인 형태소로 분해하는 과정이다. 예를 들어, "고양이가 집에 있다"라는 문장은 "고양이/명사 + 가/조사 + 집/명사 + 에/조사 + 있다/동사"로 형태소 분석을 통해 분해될 수 있다. 형태소 분석은 텍스트 분류, 감성 분석, 주제 모델링 등의 분석 작업에 선행되는 기초 작업이다.

(2) 감정 분석

감정 분석(감성 분석 또는 긍·부정 분석)은 텍스트 데이터에 담긴 긍정적, 부정적, 중립적인 감정을 파악해 제품 및 서비스에 대한 고객의 반응을 이해하는 데 사용된다. 예를 들어, 특정 제품에 대한 리뷰가 긍정적인 경우, 해당 제품의 강점을 강화하는 방향으로 마케팅 전략을 수립할 수 있으며, 반대로 부정적인 리뷰가 많을 때는 문제점을 파악하고 이를 개선하는 방향으로 전략을 조정할 수 있다.

(3) 주제 모델링

주제 모델링은 대규모 텍스트 데이터에서 주제(토픽)를 자동으로 추출하는 기법이다. 텍스트 데이터에서 자주 언급되는 주제를 찾아낼 수 있으며, 이를 통해 텍스트의 전체적인 내용을 파악할 수 있다. 예를 들어, 뉴스 기사 데이터를 분석할 때, 특정 시점에서 가장 많이 언급된 주제를 도출하고 이를 바탕으로 해당 시기의 사회적 이슈나 트렌드를 분석할 수 있다.

(4) 텍스트 분류

텍스트 분류는 텍스트 데이터를 특정 카테고리 또는 레이블로 자동 분류하는 기법이다. 텍스트 문서를 사전에 정의된 카테고리로 분류하기 위해 머신러닝 모델을 사용한다. 예를 들어, 뉴스 기사를 '정치', '경제', '스포츠' 등의 카테고리로 분류할 수 있다. 스팸 메일 필터링, 감성 분석(긍정/부정 분류), 주제 분류, 고객 리뷰 분석 등 다양한 분야에서 활용된다.

(5) 클러스터링

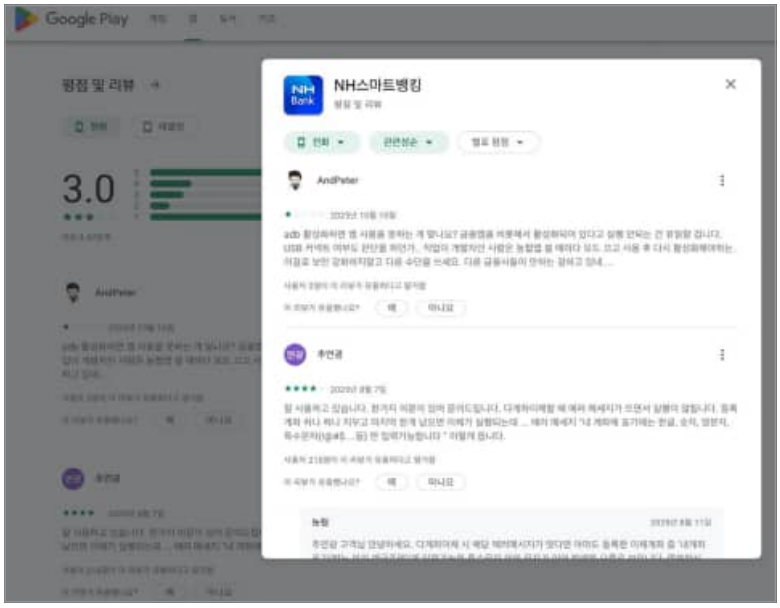
클러스터링(Clustering)은 비슷한 특성을 가진 데이터들을 그룹으로 묶어 주는 분석 기법이다. 텍스트 데이터를 클러스터링하면, 유사한 패턴이나 주제를 가진 텍스트들을 하나의 그룹으로 묶어 낼 수 있다. 이를 통해 특정 주제에 대한 전반적인 의견을 파악하거나 주요 관심사를 추출할 수 있다. 예를 들어, 호텔 리뷰 데이터를 클러스터링해 '클러스터 0 : 가성비·역세권·청결', '클러스터 1 : 객실 컨디션(공간·온도)', '클러스터 2 : 가치 극대화(조식/뷔페)' 등과 같이 클러스터별로 고객의 의견을 그룹화할 수 있다.

2. AI를 활용하여 텍스트 데이터 수집하기

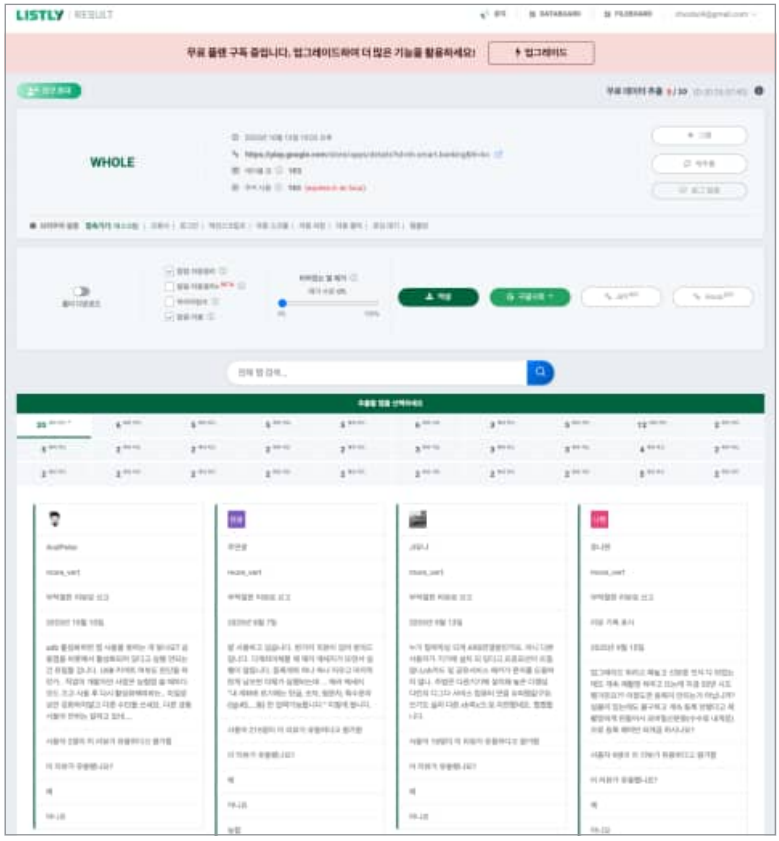
데이터 분석의 일반적인 프로세스와 마찬가지로 텍스트 데이터 분석 역시 문제 정의가 선행된다. 이 문제를 해결할 텍스트 데이터를 수집 및 전처리하고 분석 및 시각화한 뒤, 분석 결과를 활용해 새로운 가치를 만들어야 한다.

우리가 모바일 앱 서비스를 한다고 가정해 보자. 문제는 우리 서비스에 대한 고객의 불만을 개선하여 서비스의 품질을 높이는 것이라고 하자. 이 문제를 해결하기 위해서는 우리가 제공한 앱에 대한 고객의 리뷰를 앱 스토어에서 수집해야 한다. 이런 경우에 설문조사나 심층 면접 같은 방법도 있지만 고객의 진솔한 이야기를 자유롭게 적은 텍스트가 이미 존재하기 때문에 앱 혹은 웹에 있는 디지털 텍스트 데이터를 활용하는 것이 훨씬 좋은 선택이다.

예를 들어, NH스마트뱅킹 앱 이용 고객의 리뷰(Voice of Customer, VOC)를 분석하여 서비스 개선 방안을 찾아보자. 구글플레이에서 NH스마트뱅킹 앱의 리뷰를 수집할 수 있다[그림 1]. 다양한 방법이 있겠지만 여기서는 AI를 이용하여 노코드로 수집해 보자. 웹사이트의 데이터를 몇 번의 클릭만으로 수집할 수 있는 리스틀리(listly.io/ko/) 확장 프로그램을 이용한다. 수집한 데이터는 엑셀이나 구글 시트로 다운로드 가능



[그림 1] 구글플레이의 앱 리뷰



[그림 2] 리스틀리를 이용한 웹 데이터 수집

하다[그림 2].

이렇게 수집한 데이터를 작성일, 이용 후기, 유용성 평가, 답변일 등 분석에 필요한 필드만 남겨놓고 작성자 등 불필요한 필드는 제거하는 전처리를 수행한다. 여기서 한글 리뷰를 전체 복사한 다음 MS 워드에서 영문으로 기계 번역한다. 데이터 파일에 새로운 필드를 하나 생성해서 영문 리뷰를 붙여 넣기 한다. 한글 리뷰를 영문으로 번역한 이유는 텍스트 데이터 분석에서 시각화 등은 아직 한글 처리에 한계가 있기 때문이다.

인덱스	작성일	이용후기	유용성평가	답변일	Reviews
1	2025-08-07	잘 사용하고 있습니다. 한가지 의문이 있어 문의드립니다. 다계좌이체할 때 매러 메세지	30	2025-08-11	I'm using it well. I have one question, so I would like to as
2	2025-07-01	어플 너무 느립니다 처음 로그인 화면도 타행 어플에 비해 너무 느리고 로그인 버튼 누르	153	2025-07-01	The app is too slow The first login screen is also too slow
3	2025-06-26	농.축협 정기예금 모바일 가입하려고했는데 진행다해놓고 마지막 비밀번호4자리 입력하	1	2025-07-01	Cage. I tried to sign up for the livestock cooperative time
4	2025-06-25	자주쓰는 계좌(별모양 아이콘) 등록할때 띄어쓰기 공백이랑 특수문자는 등록을 못하던	2	2025-07-01	When registering a frequently used account (star icon), I c
5	2025-06-19	계좌 복붙이 안되는 유일한 은행어플 ㅋㅋㅋㅋ 복사 붙여넣기해서 어플 내 계좌번호 까	1		The only bank app that doesn't allow you to copy your ac
6	2025-07-02	카드발급하는데 맨 처음에 후불교통카드 기능 선택하는 게 있어서 선택했더니 모든 절	4	2025-07-04	When issuing a card, there was a deferred transportation
7	2025-04-07	별 1개조차 아깝습니다. 인터페이스가 최악입니다. 고객과 장난하자는 것도 아니고 정	287	2025-04-21	Even 1 star is a waste. The interface is the worst. It's not
8	2025-05-25	(추가) 업데이트 하면 해결이 될까 싶어서 업데이트 후에 otp등록을 하려고 하니 이미	32	2025-05-27	I thought that updating would solve the problem, so I tried
9	2025-05-20	최근 업데이트 뒤에... 사용 불편. 얼마 전만 해도 사용이 편했는데... 뭔 생각으로 인터	23	2025-05-21	After the recent update... Inconvenient to use. Not long ag
10	2025-04-14	아니 휴먼계정이라 그래서 휴대폰 인증하면 공지사항 페이지로 이동되고 또 로그인을	15	2025-04-16	No, it's a human account, so when you authenticate your
11	2025-05-22	그냥 원래대로해놔.. 아니 할것이없나 철쓰고있는데 쓸대없이 업데이트하고 인터페이스	15	2025-05-22	Just leave it as it was, or no, there's nothing to do, I'm usi
12	2025-05-22	오늘 갑자기 접속 인원이 많다고 대기시간 뜨면서 인증서 로그인이 안되었습니다. 그래서 혹시나 삭		2025-05-22	Today, there was a sudden large number of users, and the
13	2025-04-30	업데이트 한뒤로 비밀번호나 금액을 누를때는 잘 눌러지는데 계좌번호를 누를때 잘 안눌러짐. 계좌번		2025-05-16	After updating, when I press the password or amount, it c
14	2025-04-24	계좌번호 복사해서 붙여 넣기가 잘 안되서 너무 불편하네요..... 복사한 계좌로 이체하는거 눌러도 계		2025-04-25	It's very inconvenient because I can't copy and paste the
15	2025-08-09	로그인이 계속 안돼서 체대로 쓰지도 못하고 삭제함. 별별인증 별별 방법 이것저것 다 해봤는데 모두		2025-08-11	I couldn't log in and deleted it without being able to use it
16	2025-03-08	간편인증 로그인 안됨 -> 잠기 미거래 해제 하셔야 돼요~. 잠기미거래 해제 할려 하는데	12	2025-03-10	Easy authentication cannot be logged in -> You need to c
17	2025-02-24	업데이트 전에는 저장된 금액과 계좌이체정보로 한번 클릭으로 매달 같은 곳에 정해진	4	2025-02-25	Before the update, it was convenient to transfer a set amo
18	2025-04-14	카카오 기프티콘 결제 할려고 Ars인증 까지 하고나면 갑자기 10분동안 활동이 없어서	3	2025-04-15	After doing ARS authentication to pay for Kakao Gifticon,
19	2025-03-10	오랜만에 들어와보니 간편번호 로그인 사라져서 회원가입부터 다시하고, 모바일 otp 발아야 한다고		2025-03-14	When I came in for the first time in a long time, the simple
20	2025-03-07	디자인도 구리고 도대체 메뉴를 찾을 수가 없게 되어 있어서 쓰기 불편해요.도대체 메뉴	11	2025-03-07	The design is poor and I can't find the menu, so it's incon

[그림 3] 수집 및 전처리된 텍스트 데이터

3. AI를 활용하여 텍스트 데이터 분석하기

챗GPT와 같은 생성형 AI는 텍스트 데이터 분석에서 매우 유용하게 사용될 수 있다. 챗GPT는 자연어 처리 기반의 강력한 언어 모델로, 텍스트 데이터를 전처리하고, 감성 분석을 수행하며, 클러스터링 등의 복잡한 분석 작업을 자동화할 수 있다. 즉, 기존 수작업 분석에 비해 시간과 노력을 크게 절감할 수 있다. 특히 챗GPT에는 데이터 분석을 위한 전용 챗봇으로 "Data Analyst"가 있다.

예를 들어, 여러분은 은행 마케팅 관리자다. 어떻게 하면 텍스트 데이터에서 고객의 니즈를 파악하거나 서비스의 개선 방안을 수립할 수 있을까? 이 문제를 해결하기 위해 우리는 앞에서 리스틀리를 이용해 텍스트 데이터를 수집했다. 위에서 제시한 은행 앱 데이터를 기반으로 서비스 개선 방안을 찾아보자.

(1) 챗GPT/Data Analyst에서 텍스트 데이터 분석하기

먼저 이 데이터를 "챗GPT/Data Analyst"에 업로드하고 데이터 탐색을 시켜본다. 제공한 데이터를 챗GPT가 인식하게 하여 데이터 구조와 유형 분석 및 간단한 기술 통계분석으로 전반적인 데이터를 파악하게 된다.

분석에 앞서 프롬프트에 무엇을 요청할지 생각해야 한다. 텍스트 데이터 분석에 필요한 요청 사항들을 하나씩

작성하면 된다. 해결해야 할 문제, 즉 과제를 해결하는 데 필요한 분석이 어떤 것인지를 사전에 정리해서 차례대로 하나씩 요청하는 방안을 추천한다. 물론 이러한 정리가 어렵다면 주어진 데이터로 어떤 분석이 가능한지, 혹은 이러이러한 결과를 얻기 위해 제공한 데이터로 어떤 분석이 가능한지 챗GPT에 요청하는 것으로 분석을 시작해도 된다. 여기서는 감정 분석, 주제 모델링, 키워드 추출, 시각화 등의 분석 방법을 활용하여 앱 기능에 대한 서비스 개선 방안을 찾는다

"제공한 데이터를 탐색해 줘"라는 프롬프트에 챗GPT의 답변은 다음과 같다.

데이터셋의 기본적인 구조와 특징입니다:

□ 데이터 개요:
- 총 행(row): 460개
- 총 열(column): 6개

□ 열 설명:
- 인덱스: 리뷰 순서로 보이는 번호 (중복 없음)
- 작성일: 리뷰가 작성된 날짜 (문자열 형식)
- 이용후기: 고객이 남긴 한글 리뷰
- 유용성평가: 다른 고객이 이 리뷰를 유용하다고 평가한 수치 (일부 결측값 존재)
- 답변일: 기업의 답변 날짜 (일부 결측값 존재)
- Reviews: 영어로 번역된 리뷰 내용

□ 결측치 정보:
- 유용성평가: 85개 결측(460 - 375)
- 답변일: 156개 결측 (460 - 304)

다음 중 어떤 탐색을 원하시나요?
- 리뷰 작성일 및 활동량 분석
- 유용성 평가 분포
- 답변 유무 및 패턴
- 리뷰 내용(한글 또는 영어)에 대한 텍스트 분석
- 감정 분석 또는 키워드 추출
- 기타 직접 지정한 분석

원하는 항목을 알려주시면 그에 맞춰 탐색해드릴게요.

이제 본격적인 텍스트 데이터 분석을 해야 한다. 챗GPT가 제안한 분석 방법 혹은 문제 해결을 위해 필요하다고 생각하는 분석을 인간 분석가의 통찰을 바탕으로 진행하면 된다.

(2) 리뷰 내용에 대한 텍스트 분석

일반적으로 텍스트 분석은 텍스트 전처리(불용어 제거, 특수문자 제거 등), 형태소 분석(명사, 동사 등 주요 어휘 추출), 단어 빈도 분석(가장 많이 사용된 단어), 워드 클라우드 시각화(시각적으로 단어 중요도 표현)

등으로 이루어진다. 기존의 파이썬 코딩은 이 모든 과정을 수작업으로 진행해야 한다. AI를 활용하면 이 모든 내용을 자동으로 처리할 수 있다. 다만 아직은 한글 텍스트를 AI에서 분석하는 데 어려움이 있다.

챗GPT의 현재 환경에서는 코엔엘파이(KoNLPy, 한국어 형태소 분석기)가 설치되어 있지 않아 형태소 분석을 바로 수행할 수 없다. 다른 생성형 AI도 유사한 문제를 가지고 있다. 가장 쉬운 대안은 한글 리뷰를 영문으로 번역한 다음 영문 리뷰로 텍스트 분석을 하는 방법이다. 영문 번역은 MS 워드에서 쉽게 기계 번역을 할 수 있다. 분석의 결과는 한글로 제공된다.

"리뷰 내용(영문)에 대한 텍스트 분석을 해줘."라는 프롬프트에 대해 챗GPT가 제공한 결과는 [그림 4]와 같다.

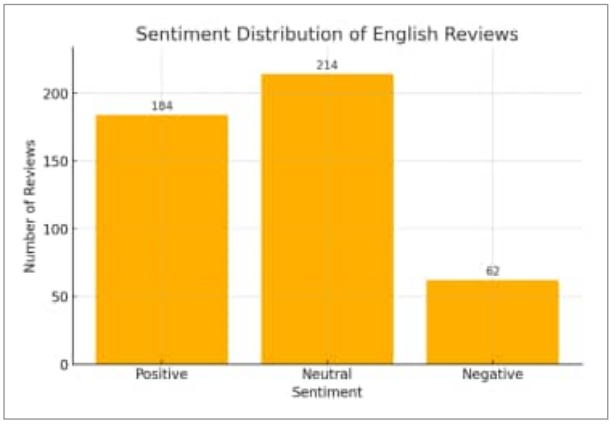


[그림 4] 리뷰 내용(영문)에 대한 워드 클라우드 맵

워드 클라우드 분석으로 영어 리뷰에서 자주 등장한 단어를 시각화한 결과다. 발견된 주요 단어는 "login, slow, error, password, transfer, screen, time, work, account, open" 등이다. 이러한 단어들을 통해 고객들이 주로 로그인 문제, 느린 속도, 이체 오류, 비밀번호 관련 이슈, 앱이 작동하지 않거나 열리지 않는 문제 등에 대해 언급하고 있음을 확인할 수 있다.

(3) 감정 분석과 키워드 추출하기

리뷰 내용에 대한 고객들의 긍/부정, 중립적 감정을 파악하기 위해 감정 분석을 요청하고, 그 결과를 시각화 차트로 분석하기 위해 "감정 분석(긍정/부정 리뷰 구분)을 해줘. 시각화 차트에 값을 표시해 줘."라고 챗GPT에 요청했다. 분석 결과는 [그림 5]와 같다. 리뷰에 대한 감정 분석 결과, 전체 리뷰 수 460개 중에서 중립(Neutral)이



[그림 5] 리뷰에 대한 감정 분석 결과

214개 46.5%, 긍정(Positive)이 184개 40.0%, 부정(Negative)이 62개 13.5%로 나타났다. 전체 리뷰 중 절반 이상이 중립적이며, 긍정 리뷰가 부정 리뷰보다 약 3배 많음을 확인할 수 있다. 앱 사용에 있어 좋은 경험을 한 고객들이 많지만, 일부는 부정적인 감정(로그인, 속도, 오류 등)을 표현하고 있다.

이 분석 결과로 보면, 중립 리뷰가 많기 때문에 고객 응대나 안내 메시지의 개선 여지가 있어 보인다. 서비스 개선을 위해서는 긍정 리뷰보다는 부정 리뷰를 살펴봐야 한다. 부정 리뷰에 자주 등장하는 단어를 추가로 추출하기 위해 "부정 리뷰에 자주 등장하는 단어를 추출해 줘."라고 요청했다.

20개의 단어가 추출됐다. 이를 바탕으로 고객이 어떤 문제를 자주 언급하는지 파악하기 위해 "부정 리뷰에서 자주 등장한 단어 상위 20개를 통해 고객이 어떤 문제를 자주 언급했는지 파악해 줘"라고 다시 요청했다. 부정 리뷰에서 자주 등장한 단어 상위 20개를 분석하여, 고객들이 자주 언급한 문제 유형을 요약한 결과는 [그림 6]과 같다.

주요 단어	관련 문제 해석	설명
transfer, account	이체 및 계좌 관련 문제	사용자는 계좌 간 이체 실패, 오류, 불편함 등을 자주 언급했습니다. 예: 다계좌 이체, 송금 실패 등
error, issue, problem, fail	앱 오류 및 작동 문제	다양한 기능에서 오류 메시지, 작동 불능 현상이 반복적으로 보고됨
login, password	로그인/인증 문제	로그인 지연, 패스워드 입력 문제, 인증 실패 등이 주요 불만 요소
time, slow, loading	느린 속도 및 대기 시간	앱 실행이나 페이지 로딩 속도가 느리다는 평가가 많음
update, version	업데이트 후 문제	업데이트 이후 앱이 더 느려졌거나, 특정 기능이 작동하지 않음
screen, open, work	앱 실행 및 UI 문제	화면이 열리지 않거나 버튼 작동이 안 되는 문제 등 인터페이스 관련 불편 호소
service, customer	고객 서비스에 대한 불만	고객센터 응답 지연, 문제 해결 미흡 등과 관련된 언급 가능성

[그림 6] 상위 단어 기반 고객 불만 유형 분석

주요 고객 불만 유형을 정리하면, 기술적 오류(앱 다운, 기능 작동 안 함), 로그인 및 인증 실패, 속도 문제(앱 지연, 느림), 이체 및 계좌 관련 오류, 업데이트 후 불안정성, 고객 서비스 및 대응 부족 등으로 나타났다. 이러한 결과를 바탕으로 앱 품질 개선, 고객 경험 개선 우선순위 결정, 고객지원 강화 등에 활용할 수 있다. 고객의 숨은 니즈를 고객이 스스로 작성한 텍스트에서 찾는 방법이다.

(4) 주제 모델링 및 토픽별 주제 분석

텍스트 데이터에서 자주 언급되는 주제를 찾아내어 텍스트의 전체적인 내용을 파악하기 위해 토픽 모델링 분석을 해보자. 단, 챗GPT의 특성 중 하나는 위와 같이 드릴다운(Drill-down, 요약된 데이터에서 시작하여 더 상세한 하위 수준의 데이터로 이동하며 분석하는 기법)이 될 때 전체 데이터로 분석 요청을 하지 않으면 제한된 데이터로만 분석이 된다는 것이다. 이는 엑셀에서 필터로 제한했을 때와 동일한 현상이다. 따라서, 분석의 목적에 따라 필요한 요청을 별도로 해야 한다. 챗GPT에게 "전체 응답을 기준으로 토픽 모델링을 수행해 줘"라고 요청했다.

총 5개의 토픽을 구했다. 각 토픽은 주제별로 자주 등장한 단어들을 기반으로 구성되어 있다. 이를 통해 고객 리뷰가 주로 어떤 주제에 집중되어 있는지 파악할 수 있다. 이후 "토픽 모델링 결과를 기준으로 고객 리뷰가 주로 어떤 주제(문제/경험)에 집중되어 있는지 파악해 줘"라고 다시 요청했다.

각 토픽에서 자주 등장한 단어들을 해석하여 고객이 리뷰에서 집중한 주요 주제(문제/경험)를 [그림 7]과 같이 요약할 수 있다.

Topic	주요 단어	주요 주제	설명
Topic 1	use, bank, banking, transfer, easy, app, account	계좌 및 이체 기능 중심의 사용 경험	계좌 관리 및 송금 기능 이용 중 편의성 또는 불편함에 대한 언급
Topic 2	app, use, account, bank, time, change, good, transaction	앱의 전반적인 품질 및 기능 변화에 대한 의견	기능 개선, 사용 시간, 전반적인 품질 평가에 대한 피드백
Topic 3	app, phone, dont, screen, use, error, mobile	앱 실행 오류 및 디바이스 관련 문제	모바일 기기에서 앱 실행 문제, 오류 메시지, 화면 멈춤 등
Topic 4	app, bank, update, apps, dont, nonghyup, just	업데이트 후 문제 및 특정 은행 관련 불만	업데이트 이후 기능 오류, 농협 관련 불만
Topic 5	log, app, time, error, im, dont, use, authentication	로그인, 인증 및 보안 이슈	로그인 실패, 인증 지연, 접근성 문제 등

[그림 7] 토픽별 고객 리뷰 주요 주제 분석

토픽 모델링을 기반으로 도출된 고객 리뷰 주제 분석 결과는 각 토픽에 대한 핵심 키워드, 주제 요약, 해석을 포함하고 있어 고객 불만 및 관심 영역을 전략적으로 파악하는 데 활용할 수 있다.

4. AI를 활용하여 서비스 개선 전략 수립하기

텍스트 데이터 분석 결과를 바탕으로 서비스 개선 방안을 챗GPT에 요청해 보자. "지금까지 분석한 결과를 바탕으로 서비스 개선 방안 5가지를 제안해 줘"라고 프롬프트에 요청했다.
리뷰 데이터 분석(텍스트 분석, 감정 분석, 토픽 모델링)을 기반으로 NH스마트뱅킹 앱의 사용자 불만과 주요 주제들을 파악하고, 이를 토대로 [그림 8]과 같은 서비스 개선 방안 5가지를 제안해 주었다.

개선 분야	주요 이슈	개선 방안
이체 기능 개선	이체 중 오류, 계좌 관련 불만	다계좌 이체 오류 원인 명확화, UI 개선, 추천 계좌 자동완성 도입
로그인 및 인증 개선	로그인 실패, 인증 복잡함	생체인식 로그인 강화, 오류 메시지 구체화, 보조 로그인 수단 제공
속도 및 반응성 향상	앱 실행 및 로딩 지연	경량화 UI 적용, 비동기 처리, 진행 상태 표시 도입
업데이트 안정성 확보	업데이트 후 기능 오류	사용자 테스트 강화, 기능 호환성 검증, 변경사항 고지
고객지원 응답 체계 고도화	문제 대응 지연, 고객 응대 부족	실시간 챗봇 도입, 리뷰 응답 확대, FAQ 지속 업데이트

[그림 8] 서비스 개선 방안 요약

고객 리뷰 분석을 바탕으로 이체 기능 개선, 로그인 및 인증 개선, 속도 및 반응성 향상, 업데이트 안정성 확보, 고객지원 응답 체계 고도화 등 서비스 개선 방안 5가지가 제안되었다. 각 항목에는 개선이 필요한 핵심 이슈, 이에 대응하기 위한 구체적 방안을 포함하고 있어, 실제 개선 전략 수립 및 우선순위 설정에 바로 활용할 수 있다.
지금까지 진행한 고객 리뷰 데이터 분석 결과를 종합하면, NH스마트뱅킹 앱은 핵심 금융 기능의 품질 안정성과 사용자 경험 전반에서 개선이 필요한 지점을 명확히 보여주고 있다. 제안한 5가지 개선 방안을

실행하면 고객 만족도와 앱 신뢰도를 동시에 향상할 수 있을 것이다. 다만, 여기 제안된 내용은 특정 기업의 특정 서비스에 국한된 내용이기 때문에 사례로서는 의미가 있지만 일반화할 수는 없다. 텍스트 데이터 수집 및 분석 방법은 예시한 방법을 그대로 활용하면 된다.

5. AI를 활용한 텍스트 데이터 분석 시사점

여러분은 여러분의 문제를 정의하고 텍스트 데이터를 수집하고 분석하여 필요한 결과물을 얻어야 한다. AI가 분석하고 제안한 내용에 대해 좋다거나 나쁘다거나 판단을 내리는 것은 무의미하다. 데이터 분석에 대한 기본적인 이해를 바탕으로 문제를 해결하기 위해 분석적 마인드를 가지고 심층적인 분석을 해야만 원하는 결과, 즉 문제를 해결할 수 있는 분석 결과를 얻을 수 있다.

AI를 활용하지 않고 단기간에 텍스트 데이터를 손쉽게 수집, 분석하며 개선 방안을 찾는 것은 현실적으로 매우 어려운 일이다. 다만 분석 결과와 제안 내용을 그대로 수용하는 것은 문제의 소지가 있다. 분석 자체는 객관적이지만 그 과정에 주관이 개입되기 때문이다. 분석가의 주관이든 AI의 주관이든 분석 결과는 주관적인 결과물이다. 단지 데이터로 분석했기 때문에 객관적으로 보일 뿐이다.

따라서 AI가 분석하고 제안한 결과를 바탕으로 인간 분석가의 통찰이 더해져야 한다. AI는 책임을 지지 않기 때문이다. 분석 결과는 문제 해결과 의사결정을 위한 기초 자료이며 이를 옥석으로 만드는 것은 인간 분석가의 몫이다. 도메인(예시한 사례에서 은행 업무 및 앱 서비스)에 대한 지식이 있어야 데이터 분석의 결과를 보다 의미 있게 통찰할 수 있다. 여전히 인간 분석가가 필요한 이유이다.

참고문헌

구자룡(2025).『챗GPT로 시작하는 데이터 리터러시』. 마들렌북.

구자룡(2024).『AI 데이터 분석』. 커뮤니케이션북스.

구자룡(2024).『데이터 마인드 기르는 습관』. 좋은습관연구소.



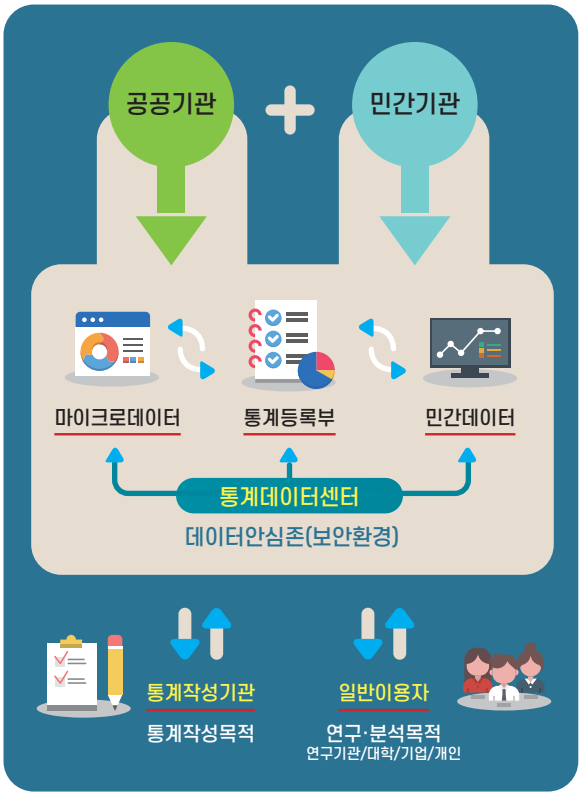
SDC
Statistics
Data Center

행정통계자료와 민간자료를 한곳에!

통계데이터센터 서비스

통계데이터센터가 새로운 서비스로 정보화 사회를 선도합니다.

국가데이터처가 제공하는 통계등록부, 마이크로데이터, 민간데이터 등 다양한 자료를
한 곳에서 연계 및 분석이 가능한 통계데이터센터(SDC)



1 분석 플랫폼 제공 서비스

- 분석시스템 · 통계패키지 제공
- 통계자료(통계등록부 · 통계기초자료) 및 민간자료, 이용자 반입자료 연계 · 분석

2 전문가 분석지원 서비스

- 분석 경험이 없는 이용자를 위한 데이터 분석 지원
- 센터 이용 상담 및 데이터 분석 자문

3 주문형 분석서비스

- 시간 및 거리상 센터 방문이 어렵거나 직접 자료분석을 하기 힘든 이용자를 위한 서비스
- 센터 이용자료를 활용하여 연계 · 분석 후 이용자가 원하는 형태로 결과를 제공

4 명부 서비스(승인통계대상)

- 분석센터로 방문하여 자료분석 및 표본설계를 통해 데이터 반출

5 이용자 교육 서비스

- 이용자 교육 홈페이지 운영
- 통계분석 프로그램 및 분석사례 교육
- 매년 통계데이터 활용대회 개최

국가데이터처, 정부부처, 지방자치단체, 연구기관 등

모든 기관의 마이크로데이터를 한 곳으로



보다 심도 있고 다양한 분석을 원한다면
지금 바로 MDIS를 클릭해 보세요.

■ 서비스 소개 (2025년 10월 기준)

가. 서비스명 : 마이크로데이터통합서비스(MDIS, mdis.kostat.go.kr)
나. 제공 통계 수 : 174개 작성기관 통계 총 401종 제공(국가데이터처 50종)
다. 제공 형태 : 마이크로데이터(조사대상에 따라 인구, 사업체, 가구 기반 자료)

기준		통계명
국가 데이터처 (50종)	인구·가구 (15종)	가계금융복지조사, 가계동향조사, 가구소비실태조사, 경제활동인구조사, 국내인구이동통계, 녹색생활조사, 사망원인통계, 사회조사, 생활시간조사, 이민자 체류실태 및 고용조사, 인구동향조사, 인구총조사, 주택총조사, 지역별고용조사, 초중고 사교육비조사
	사업체 (12종)	건설업조사, 경제총조사, 광업제조업조사, 기업활동조사, 도소매업조사, 서비스업조사, 서비스업총조사, 소상공인실태조사, 운수업조사, 전국사업체조사, 전문과학기술서비스업조사, 프랜차이즈조사
	농림어업 (12종)	농가경제조사, 농림어업조사, 농림어업총조사, 농산물생산비조사, 농어업법인조사, 농업면적조사, 농작물생산조사, 세종시특별센서스, 양곡소비량조사, 어가경제조사, 어류양식동향조사, 어업생산동향조사
	행정통계 (11종)	귀농어 · 귀촌인통계, 기업생멸행정통계, 생애단계별행정통계, 신혼부부통계, 연금통계, 영리법인기업체행정통계, 육아휴직통계, 일자리이동통계, 일자리행정통계, 임금근로일자리동향행정통계, 주택소유통계
통계작성기관 (351종)		전국다문화가족실태조사, 가족실태조사, 노인실태조사, 자동차주행거리통계, 직종별사업체노동력조사, 국민여가활동조사, 외래관광객조사, 가공식품소비자태도조사 외 343종

■ 서비스 내용

가. 구분 : 자료의 민감성 정도에 따라 공공용, 인가용으로 구분 운영

나. 서비스 방법

- 공공용
  - 다운로드 서비스 : MDIS 포털에서 직접 무료로 다운로드
  - 온라인분석 서비스 : MDIS 포털 내 분석 시스템에서 무료로 추출 · 편집 · 분석
- 인가용
  - 원격접근 서비스 : 승인 후 이용자가 집 · 사무실 등에서 국가데이터처 서버 접속 후 활용
  - 센터 서비스 : 승인 후 이용자가 분석센터에 방문하여 활용
  - 주문형 서비스 : 승인 후 이용자에게 맞춤형으로 제공

다. 수수료

- 공공용 : 무료
- 인가용 : 유료

■ 문의

- 연락처 : (재) 한국통계진흥원
- 전화 : (02) 3457-9700, FAX :(02) 515-0240
- 주소 : (06097) 서울특별시 강남구 선릉로 612, 6층
- E-mail : MDIS@stat.or.kr, mdis.kspi@gmail.com

국가데이터처에서 국가통계를 활용하세요!

국가데이터처는 통계개발·활용·교육에 필요한 모든 정보와 도움을 제공합니다.
다양한 국가통계정보 제공 사이트를 활용하세요.











통계인재개발원

shi.kostat.go.kr

국내 유일의 국가통계교육 전문기관

통계 작성 및 활용 전문통계과정,
기관맞춤형과정, e-러닝 과정

통계데이터센터

data.mods.go.kr

행정통계자료와 민간자료를 한곳에

행정통계자료(통계등록부), 민간자료의
연계·융합이 가능한 데이터 플랫폼

MDIS

mdis.kostat.go.kr

원하는 자료를 직접 분석 및 요청

다운로드 서비스/온라인분석 서비스 선택 시
공공용 마이크로데이터를 무료로 분석 활용 가능

KOSIS

kosis.kr

국가통계 쉽게 찾기

국내, 국제, 북한의 주요 통계를
한 곳에 모아 알기 쉽게 분류해 제공

SGIS

sgis.mods.go.kr

지도 위 통계정보 살펴보기

인구, 가구, 주택, 사업체 통계 등 각종 통계를
지도(GIS) 위에서 한눈에 파악

국가데이터처
통계인재개발원