통계의 창

WINDOW OF STATISTICS

2025 SUMMER
Vol. 35



인구센서스 100년

센서스 100년, 2025 인구주택총조사 실시

인구주택총조사 100년 모아보니 '대한민국 발전史'였다

지속가능한 사회를 위한 포용적 센서스를 위한 제언

인구주택총조사 100년: 센서스의 의의와 향후 발전방안

CONTENTS

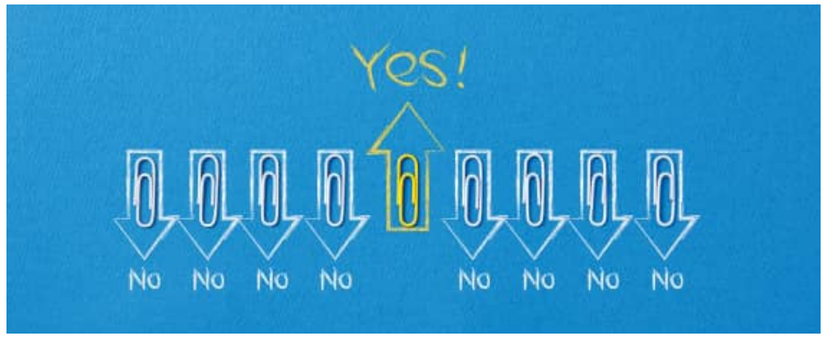
통계의 창
2025 SUMMER
Vol. 35

발행일 2025년 6월 2일
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

2025 Census 센서스 100년



통계 ISSUE

4 **센서스 100년, 2025 인구주택총조사 실시**
정은숙 | 인구총조사과 서기관

12 **인구주택총조사 100년 모아보니 '대한민국 발전史'였다**
박소정 | 조선비즈 기자

18 **지속가능한 사회를 위한 포용적 센서스를 위한 제언**
박효민 | 서울시립대학교 도시사회학과 교수

22 **인구주택총조사 100년: 센서스의 의의와 향후 발전방안**
전광희 | 충남대학교 사회학과 명예교수

※ '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.

통계 이야기

28 **AI가 바꾸는 교실: 새로운 리터러시의 시대**
김용성 | 충남대학교 기술교육과 교수

32 **학교 통계 교육에서 데이터 윤리**
권나영 | 인하대학교 수학교육과 교수

36 **생성형 AI 시대의 필수 역량, 디지털 리터러시**
김한성 | 고려사이버대학교 소프트웨어교육 교수

44 **통계로 바라보는 세상이야기**
신동헌 | 도서출판 지일북스 대표

48 **데이터 앤 의사결정 리터러시 연결**
강양석 | DEEP SKILL 대표

58 **AI를 활용하여 서베이 및 고객 데이터 분석하기**
구자룡 | 밸류바인 대표

ISSUE 01



센서스 100년, 2025 인구주택총조사 실시

정은숙 | 인구총조사과 서기관



인구주택총조사 의의

인구주택총조사는, 특정한 시점에 한 국가의 모든 사람, 가구, 거처와 관련된 인구 · 경제 · 사회 자료를 수집, 집계, 분석, 제공하는 전체 과정이라고 유엔(UN: United Nations)에서 정의하고 있다.

인구주택총조사, 센서스는 오랜 역사를 가지고 전 세계적으로 실시되고 있는 통계조사이다. 근대적인 센서스는 미국에서 1790년에 최초로 실시하였고, 영국은 1801년, 캐나다는 1871년, 우리나라는 1925년 부터 실시하고 있다. 2020 센서스 라운드('15~'24년) 동안 239개 중 204개[1] 국가(또는 지역)가 최소 1회 이상의 센서스를 실시하였다.

센서스는 국가 영토 내의 모든 사람과 거처(완전성)를, 개별적(개별성)으로, 기준시점에서 일정 기간(동시성) 동안, 일정한 주기(주기성)로 파악하는 국가 기본 통계조사로, 읍면동 단위까지 통계가 작성 · 제공되고 있다.



미국, 1790년 센서스의 조사원 기록자료[2]



한국, 1925년 간이국세조사 조사표

1) 출처: 제56차 유엔통계위원회 보고서('25.3월)
2) 출처: US Census Bureau 홈페이지

인구주택총조사 활용

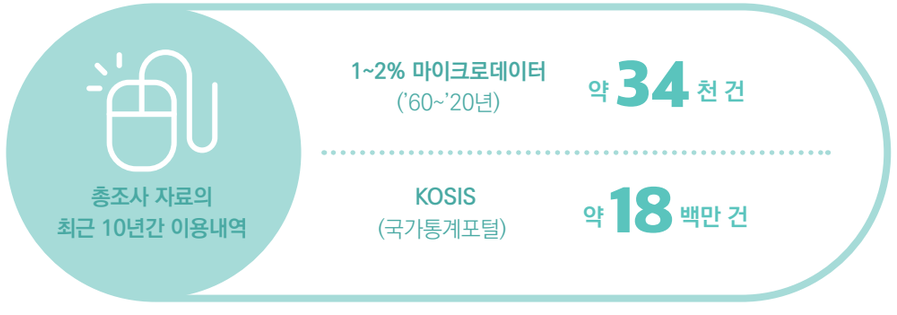
인구주택총조사의 결과는 중앙행정기관과 지방자치단체 등 국가 주요 정책 수립을 위한 기초자료로 활용되고 있다.

중앙행정기관 및 지방자치단체의 주요 활용 사례

활용기관	주요정책 및 사업
행정안전부	지방자치단체 외국인 주민 현황
국토교통부	주거 종합 계획, 주택보급률, 주거실태조사, 최저 주거 기준 미달 가구 지수
기획재정부	주택시장 안정 대책, 인구 · 가구 구조변화 대응 정책
보건복지부	아동 · 노인 · 미혼부모 · 장애인 정책, 각종 정책 재정 추계
교육부	평생 교육 진흥 기본계획
산업통상자원부	에너지 기본계획
여성가족부	양성 평등 정책 기본계획, 한부모 · 다문화가구 · 이주배경아동청소년 정책
고용노동부	중장기 인력 수급 전망, 노동력 수급 계획
문화체육관광부	관광 개발 기본계획
지방자치단체	지역 경제 활성화 계획, 도시 기본계획

또한 국민의 사회 · 경제적 특성을 조사하는 가구 부문 표본조사의 모집단[3] 역할을 수행하고 있고, 통계청의 장래 인구 · 가구 추계, 국토교통부의 주택보급률, 최저 주거 기준 미달 가구 등 2차 가공통계 작성에도 활용되고 있다.

그 외 대학 및 연구기관의 연구, 민간 기업체의 경영 기초자료로 제공되어 새로운 부가가치 창출 및 상품 기획 등에 활용되고 있다. '15년부터 무상 제공하고 있는 1~2% 마이크로데이터('60~'20년)는 최근 10년간 34천 건 넘게 이용되고 있고, KOSIS(국가통계포털) 자료는 최근 10년간 18백만 건 이용으로 통계청 자료 중에 가장 많이 활용되고 있다.



우리나라 인구주택총조사 연혁

연도별 인구주택총조사 주요 현황

조사기준일	명칭	주요 특징
1925. 10. 1.	간이국세조사	인구총조사 최초 실시
1960. 12. 1.	인구주택국세조사	주택총조사 최초 실시
1966. 10. 1.	인구센서스	10% 표본조사 병행
2005. 11. 1.	인구주택총조사	인터넷조사(PC) 최초 실시
2015. 11. 1.	인구주택총조사	전수조사는 등록센서스 도입, 표본조사는 20% 확대 실시
2020. 11. 1.	인구주택총조사	표본조사는 GIS 기반 CAPI, 모바일 및 전화 조사 도입

우리나라 인구총조사는 1925년부터, 주택총조사는 1960년부터 5년마다 실시하고 있으며, 대한민국 정부 수립 이후 최초의 총조사는 1949년에 실시하였다.

100년의 총조사 중 가장 큰 변화는 2015년 결합(Combined) 센서스 방식 도입이였다. 모든 국민을 대상으로 현장조사를 실시했었던 전수조사를, 행정자료를 활용한 등록센서스로 전환하고, 행정자료로 파악되지 않는 특성항목 조사를 위해 표본 20%를 대상으로 현장조사를 실시하게 되었다.

결합(Combined) 센서스 방식

구분	조사항목	실시 주기	조사방법
전수조사	기본항목[4]	매년	행정자료 활용
표본조사	특성항목[5]	5년(0, 5자 연도)	인터넷 · 전화 · 방문면접 등 현장조사 실시

2025 인구주택총조사(표본) 준비

올해 실시하는 인구주택총조사는 센서스 100년이 되는 의미있는 해이며, 통계청은 성공적인 2025 인구주택총조사 실시를 위해 체계적으로 준비하고 있다.

조사항목, 업무 개선, 홍보, 현장조사 효율화를 위한 연구용역과 컨설팅을 추진하였고, 특히 조사항목 선정을 위해, '21년 연구용역을 시작으로, 다양한 이해관계자 의견 수렴과 전문가 자문, 유관부처와의 회의 등 총 48회의 절차를 거쳤다.

노후화된 센서스 시스템은, '21년 ISP를 시작으로 컨설팅을 거쳐, '22년부터 차세대 센서스 통합관리 시스템을 구축하여 최종 점검하고 있다.

3) 통계청 통계정책관리시스템의 '25.2월 승인통계 기준, 가구부문 표본조사의 69.2%가 모집단으로 활용
4) 성별, (만)나이, 가구주와의관계, 국적, 국적취득연도, 입국연원, 1년전거주지, 가구구분, 주거시설형태, 거처의종류 등
5) 교육정도 · 영역, 아동보육, 활동제약(돌봄), 통근통학, 경제활동상태, 혼인상태, 방의 수, 난방시설, 주차장소, 거주층 등

또한, ’22년부터 ’24년까지 3차례의 시험조사를 통해, 신규 조사항목, 차세대 센서스 통합관리시스템과 다양한 조사방법별 기능들을 테스트하였고, 작년에 실시한 시범예행조사를 거쳐 총조사 전체 과정을 점검 완료하였다.

총조사 추진전략별 중점 추진사항을 담은, 2025 인구주택총조사 기본계획안은 작년말 통계위원회[6] 본회의 심의를 거쳐 최종 확정되었다.

센서스 100년, 2025 인구주택총조사(표본) 실시

이번 2025 인구주택총조사에서는 국민의 총조사 참여 확대를 위해 응답 접근성과 편의성을 높이고, 총조사 자료 활용성 제고를 위해 다변화 사회를 반영한 신규 통계를 생산하고, 총조사 효율성과 정확성 제고를 위해 총조사 전 과정을 디지털화하려고 한다.

이를 위해, 첫째, 총조사 응답 접근성과 편의성을 높여, 국민의 적극적인 총조사 참여를 확대하고자 한다.

국민이 언제 어디서나 원하는 방식으로 응답할 수 있도록, 사전에 표본가구로 우편 발송하는 조사 안내문을 통해 다양한 조사방법별 참여 방식을 알리고, 삽입된 QR코드로 모바일 조사표에 바로 접속할 수 있도록 하였다. 전주기 자가응답률이 가장 높았던 모바일조사표를 더욱 개선하여, 휴대폰의 작은 화면에서도 최대한 편하게 응답할 수 있도록 모바일 환경에 최적화된 조사표를 설계하고 있다.

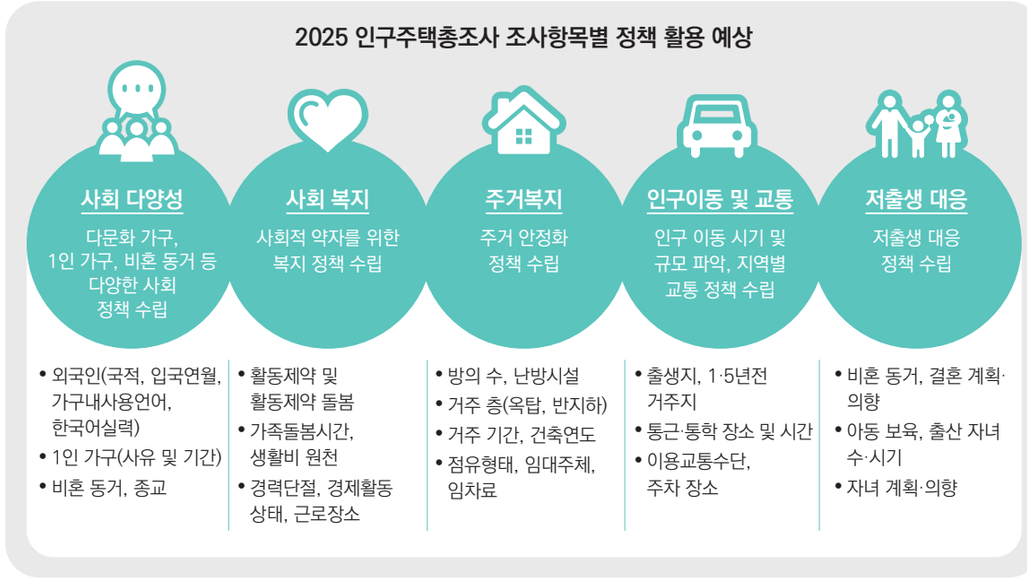
응답자가 전화조사, 대면조사를 희망하는 경우 원하는 시간으로 예약할 수 있으며, 조사기간 동안 국민의 궁금증을 24시간 신속하게 해결할 수 있도록 AI[7] 기반의 콜센터를 운영할 계획이다.

최근 증가하는 다양한 국적의 외국인 응답자를 위해, 외국어 조사표를 전주기 10종에서 20종[8]으로 확대하고, 국민의 적극적인 참여를 위해 성별, 연령, 가구 유형 등을 고려한 세분화된 홍보를 실시하여, 총조사 이해도를 높이려고 한다.

2025 인구주택총조사 실시 개요

실시근거 통계법 제5조의4 및 시행령 제2조의4, 통계법 제17조 제1항, 인구주택총조사 규칙
조사기준시점 ’25. 11. 1. 0시 현재
조사기간 ’25. 10. 22. ~ 11. 18. (28일간)
- 인터넷 · 전화조사 사전 실시: 10. 22. ~ 10. 31.
- 방문면접조사 실시: 11. 1. ~ 11. 18. ※ 기간 중 인터넷 · 전화조사 참여 가능

조사대상 조사기준시점 현재, 20% 표본조사구 내 모든 내 · 외국인 및 거처
조사항목 총 55개 중 현장조사 42개, 행정자료 등으로 대체 13개
조사방법 인터넷(PC, 모바일) · 전화 · 태블릿PC 조사, 종이조사표 활용
조사체계 통계청(주관기관), 지방자치단체(실시기관)
조사인력 30,000명 내외



둘째, 다양한 사회 변화상을 반영할 수 있도록 조사항목을 개선 개발하여, 총조사의 자료 활용도를 높이고자 한다.

2025 총조사 조사항목은, 총 55개이며, 이중 현장조사는 42개, 행정자료 등으로의 대체는 13개이다. 조사항목 총수는 전주기와 동일하나, 행정자료 등으로의 대체[9]를 확대하여 국민의 응답 부담과 현장조사 부담을 경감하려고 한다. 특히 전주기 민감항목이였던 출산자녀수와 자녀출산시기는 행정자료로 대체된다.

연도별 조사항목 현황

구분	‘00년	‘05년	‘10년	‘15년	‘20년	‘25년
합계	50	44	50	52	55	55
현장조사	50	44	50	49	45	42
행정자료 등 대체	-	-	-	3	10	13

6) 통계법 제5조의2에 의해 설치된 위원회(위원장: 부총리)로 통계의 작성 · 보급 및 이용에 관한 사항 심의 · 의결
7) 보이스봇(말로하는/누르는ARS), 챗봇, STT(Speech To Text)
8) ‘20년(10종): 영어, 중국어, 베트남어, 러시아어, 인도네시아어, 태국어, 몽골어, 일본어, 캄보디아어(크메르어), 네팔어
’25년(20종): ‘20년 10종 + 우즈베크어, 카자흐어, 미얀마어, 스리랑카어(싱할라어), 방글라데시어(벵골어), 파키스탄어(우르두어), 키르기스어, 말레이어, 우크라이나어, 라오스어
9) ’20년(10개): 1 · 5년 전 거주지, 주거시설 형태 · 수, 타지주택소유여부, 거처의종류, 주거용연면적, 대지면적, 건축연도, 총방수
’25년(13개): ‘20년 10개 + 출산자녀수, 자녀출산시기, 가구구분(가구원수, 가구주와관계 입력값으로 대체)

가족돌봄시간, 종교[10년주기], 비혼동거, 결혼 계획 · 의향, 가구내 사용 언어[이주배경가구], 한국어실력[외국인], 임대주체 등이 신규 추가되고, 반려동물, 소방시설보유여부 등 타 조사와 중복되는 항목은 조사 중지되며, 응답 부담이 높아 정확한 조사가 어려웠던 혼인연월(초혼기준), 사망자녀 수는 조사하지 않는다.

전주기 대비 조사항목 변경사항

구분	주요 내용
신규(7)	가구내사용언어, 한국어실력, 종교, 결혼계획 · 의향, 가족돌봄시간, 자전거보유, 임대주체
중지(7)	혼인연월, 결혼전취업여부, 사회활동, 반려동물, 마시는물, 소방시설보유여부, 주거전용 · 영업겸용여부
변경(18)	가구주와의관계(비혼동거 범주 추가), 활동제약(조사대상 · 척도 확대), 출산자녀수(사망자녀 수 삭제), 자녀계획 · 의향(조사대상 확대) 등

2025 인구주택총조사 조사항목

구분	표본 항목(55)			
	현장조사(42)			행정자료 등(13)
인구 (35)	①성명 ②성별 ③생년월일(나이) ④가구주와의 관계 ⑤국적 ⑥입국 연월 ⑦가구내 사용언어 ⑧한국어 실력 ⑨출생지 ⑩종교 ⑪교육 정도	⑫교육 영역 ⑬아동 보육 ⑭활동 제약 ⑮활동 제약 돌봄 ⑯통근·통학 여부 ⑰통근·통학 장소 ⑱이동교통수단 ⑲통근·통학 소요시간 ⑳경제활동 상태 ㉑종사상지위 ㉒산업	㉓근로장소 ㉔직업 ㉕현 직업 근무연수 ㉖혼인상태 ㉗결혼 계획· 의향 ㉘자녀 계획·의향 ㉙경력단절 ㉚가족돌봄시간 ㉛생활비 원천	①1년 전 거주지 ②5년 전 거주지 ③출산 자녀 수 ④자녀 출산 시기
가구 (14)	①1인 가구 사유 ②혼자 산 기간 ③방의 수 ④난방 시설	⑤자전거 보유 ⑥자동차 보유 및 주차장소 ⑦건물 및 거주층 ⑧거주 기간	⑨점유 형태 ⑩임대 주체 ⑪임차료	①가구구분* ②주거시설형태 ③타지 주택 소유 여부
주택 (6)				①거처의 종류 ②주거용 연면적 ③대지면적 ④건축연도 ⑤총방수 ⑥주거시설수

* 가구원 수 및 가구주와의 관계 입력값으로 대체

셋째, 최신의 정보통신기술을 활용하여 총조사의 정확성과 효율성을 높이고자 한다.
AI 기술을 활용하여 산업, 직업 응답 내용을 표준분류 기준으로 자동 분류하여 자료처리 업무량을 감축하고 품질을 제고할 예정이며, 조사원이 태블릿PC를 통해 현장에서 오류를 즉시 수정할 수 있도록 내검 기능[10]도 강화하려고 한다.
현장조사부터 자료 공표 단계까지 총조사 전 과정을 차세대 센서스 통합관리시스템으로 구축 운영하여, 체계적이고 안정적으로 총조사를 실시할 예정이다.

2025 인구주택총조사 참여 필요

2025 인구주택총조사는, 오는 10월 22일부터 모바일 또는 PC를 활용한 인터넷조사와 콜센터(080-2025-2025, 08~21시) 또는 시군구 상황실을 통한 전화조사가 먼저 실시된다. 미완료 가구를 대상으로 11월 1일부터 18일까지 조사원이 태블릿PC를 활용하여 방문 면접 조사를 하게 되며, 조사원의 가구 방문 조사 기간 중에도 인터넷 또는 전화조사로 응답할 수 있다.
인터넷 · 전화조사 사전 실시 1~3일 전에, 표본가구를 대상으로 총조사 대상 가구임을 알리는 조사안내문이 우편 발송된다. 표본가구 주소 정보만 있어, 해당 가구에 사는 분 이름을 특정할 수가 없기 때문에, 우편물 수신자는 가구주(또는 거주자) 귀하 등으로 명시할 예정이다. 조사안내문에는 총조사 안내, 참여 방법, 자가응답을 위한 참여번호 등이 기재되어 있어, 누구나 쉽게 다양한 조사방법으로 총조사에 참여할 수 있다.
언제나 그러하듯, 인구주택총조사의 성공적 수행을 위해서는 국민들의 적극적인 참여와 정확한 응답이 절실하다. 통계청은, 인구주택총조사 결과가 국민에게 필요한 바른 정책으로 되돌아갈 수 있도록 최선을 다할 것이며, 센서스 100년이 되는 2025 인구주택총조사에도 국민의 적극적인 협조와 많은 관심을 부탁드린다.



10) '20년 7종 → '25년 40종

ISSUE 02

인구주택총조사 100년 모아보니 '대한민국 발전史'였다

박소정 | 조선비즈 기자



2025년 '인구주택총조사'의 해가 돌아왔다.
센서스로도 불리는 이 조사는 통계청이 5년에 한 번 전 국민을 대상으로 실시하는
우리나라에서 가장 규모가 큰 국가통계조사다. 센서스를 통해 얻어지는 통계는 인구·가구·주택 등
우리 삶에 관한 극히 '기초적'인 정보들이지만, 동시에 그 시대의 사회·경제를 파악하기 위한
가장 '중요한' 기록 자료다. 그런 센서스가 시행된 지 어느새 100년이 흘렀다.
100년 동안 쌓인 센서스가 보여주는 대한민국의 역사는 어떤 모습일까.

센서스 결과가 보여주는 대한민국 변천사

1925~40 조선총독부 조사가 '근대적 센서스'의 시초

1925년 10월 1일 0시 우리나라 첫 '센서스(총조사)'가 집계됐다. 당시 총인구수는 1952만2945명, 남자와 여자는 각각 1002만943명, 950만2002명. 오늘날과 비교해 절반 이하인 인구수, 여자보다 남자가 더 많았던 성비 등은 생소한 그때의 시대상을 보여준다. 역사적으로 진귀한 가치가 있는 한국 센서스의 효시이지만, 이것이 작성된 사연은 그리 달갑지만은 않다.

최초의 센서스는 일제강점기 조선총독부가 실시했다. 당시 '간이 국세조사(國勢調査)'라는 이름으로 조사가 이뤄졌다. 인구통계는 '통치'의 기본이 된다. 일제가 조선의 형세를 파악해 노동력을 착취하고 경제적으로 수탈하는 데 활용했을 가능성이 크다.

1925년 조사를 근대적 센서스의 효시로 보는 이유에 대해 통계청 관계자는 "유엔(UN)이 규정한 센서스 원칙 권고안에는 '영토 내 모든 사람을 조사해야 한다(보편성)' '일정 시점을 기준으로 한다(동시성)' 등의 기준이 있다"며 "(간이 국세조사는) 이런 기본 원칙을 상당 부분 따른 형태"라고 했다. 그러면서 "고려시대나 조선시대에 행해졌던 유사한 형태의 인구조사는 노비나 여자, 노인은 세지 않는 등 사람을 모두 포괄하지 않았기 때문에 근대적 의미의 센서스로 볼 수 없다"고 했다.

1944~55 분단 후 인구 '뚝'… 전쟁통 한쪽 다리 잃은 男 많아



100년 전 시작된 센서스는 이후 5년마다 시행돼 왔다. 그런데 1940년대의 조사는 5년 단위로 똑떨어지지 않는 점이 관찰된다. 1945·1950년 시행됐어야 할 조사가 1944년과 1949년에 이뤄진 것이다. 이는 전쟁·해방의 역사와 관계 깊다. 중일전쟁 중이던 1944년엔 인력·물자 동원을 위해, 대한민국 정부 수립 직후인 1949년엔 국가를 제대로 운영하기 위해 1년씩 당겨 시행한 것으로 추정된다. 1949년부터는 센서스 시행 주체도 우리 통계청의 전신인 공보처 통계국으로 바뀐다.

센서스 기록은 남북 분단의 현실도 보여준다. 1944년 2590만 명까지 지속해서 증가하던 인구수는 1949년 2018만9000명으로 고꾸라진다. 1945년 8·15 광복 이후 한반도가 남북한으로 갈리게 된 기점을 전후로 한 변화다. 1944년 센서스가 남북한 인구를 합산한 마지막 통계인 셈이다. 다만 인구수가 절반으로 급감하지는 않는데, 해방 직후 인구 이동이 굉장히 잦았던 시기라는 특수성이 반영된 결과로 전문가는 추정한다.

1955년 센서스엔 한국전쟁(6·25전쟁)의 상흔도 남아 있다. 당시 인구의 1%가량인 18만 명은 '불구자(不具者·몸의 어느 부분이 온전하지 못한 사람·당시 표기)'였고, 이들의 70%는 남자였다. 남녀 모두 '농아(聾啞·듣지 못하거나 말하지 못하는 사람)'가 가장 많았고, 남자는 한쪽 다리가 잘리거나('족절단-단') 한쪽 눈이 보이지 않는('실명-단') 경우가 그다음으로 많았다.

1960~95 출산 급증…아파트가 주요 주택 형태로 부상

1960년대 들어서 정부는 센서스 결과를 본격적으로 국가 정책에 활용하기 시작했다. 당시는 1차 베이비붐 세대가 막 태어나기 시작하던 시기로, 전후(戰後) 출산율 급증이 사회 문제로 부상하던 때다.

이전까지 5년 새 6%가량 늘던 총인구수는 1955~1960년에 16% 넘게 증가했다. 1962년 공표된 '제1차 경제개발오개년계획'에는 이런 문제를 해결하고자 '가족계획(산아제한)' 정책을 포함했는데, 여기엔 '인구총조사 자료를 활용한다'는 표현이 적혀 있다.

자연스레 이 시기 미래 인구 추계도 시작됐다. 센서스 인구를 기초로 출생·사망, 인구 이동을 고려해 미래의 인구를 예측하는 것을 뜻한다. 정부의 첫 공식 인구 추계로 1964년 발표된 '한국의 신인구 추계(1960~2000년)'에 따르면, 당시 인구 정점은 1980년(4193만 명)으로 추정된 바 있다.

산업화가 빠르게 진행된 1970년대 센서스는 본격적으로 가구의 '부(富)'를 측정하기 시작했다. 1970년 가정집(전체 557만6000가구)에서 가장 많이 보유하고 있는 가전·전자기기는 라디오(404만3000가구)와 재봉틀(244만6000가구)로 집계됐다. 피아노 및 올갠(오르간·6만1000가구), 냉장고(12만7000가구)가 있는 집은 드물었다. 통계청 관계자는 "국민의 문화생활 양상 내지는 중산층 규모를 측정하는 척도가 됐다"고 설명했다.

1980~1990년대엔 급격한 도시화와 인구 집중 문제가 대두됐다. 1985년 도시별 1㎢당 인구밀도는 서울시가 1만5921.1명으로 압도적으로 높았고, 그 뒤를 부산시(8088.8명)·인천시(6713명)·대구시(4460.6명)가 이었다. 오늘날 부산·대구 인구 밀도는 그 당시와 비교해 반토막 났는데, 반면 경기도는 그때와 비교해 세 배나 높아졌다.

아파트가 우후죽순 생겨나기 시작하던 1990년대엔 주택(716만386채)의 23%(162만8117채)가 아파트였다. 여전히 단독주택(66%·472만6933채) 형태가 대부분이었지만, 아파트는 우리나라에서 두 번째로 많은 유형의 주택으로 자리 잡았다. 물론 지금은 아파트 비중(64%)이 가장 높다.



1960 대한민국 인구 및 주택 국세조사표

2000~25 기혼 女 93% "자녀 더 안 낳아"… 3집 중 1집 '혼자 산다'

2000년 한국은 'Y2K(컴퓨터 2000년 연도 인식 오류)' 문제가 전국적으로 우려를 불러일으켰듯, 개인용 컴퓨터(PC) 보급 속도가 빨랐다. 전체 가구(1431만126가구) 중 21.7%(310만9559가구)가 인터넷 회선을 보유하고 있는 것으로 집계됐다.

2005년부턴 본격적으로 저출생 문제가 불거지기 시작했다. 15세 이상 기혼 여성 1441만3990명 중 93%(1334만3585명)가 '추가 자녀 계획이 없다'고 대답했다. 사회의 개인화도 급속도로 이뤄졌는데, 1975년 4.2%에 불과하던 1인 가구 비율은 2020년 세 집 중 한 집꼴(31.7%)로 비중을 크게 늘렸다.

센서스 질문 항목이 보여주는 그 시대 관심사

센서스를 구성하는 질문 항목은 그 시대의 '관심사'를 대표한다. 1925년 최초의 센서스에선 성명, 성별, 출생연월일(당시는 출생연월), 배우 관계(혼인 상태), 민적·국적(당시는 본적·국적) 등 단 다섯 개 항목만 조사했다. 이는 인구의 기본 항목으로 현재까지도 명맥을 유지하고 있다. 100년 전 다섯 개로 시작했던 센서스 항목 수는 오늘날 55개에 이른다.

'手足 절단' '문맹 여부' 묻던 일제강점기·한국전쟁 이후

일제강점기에 센서스는 전시 동원 인력을 파악하기 위한 목적으로 수집됐다. 중일전쟁 중 실시된 1940·1944년 센서스가 그렇다. 1940년엔 '지정(指定·행정 관청이 자격을 준) 기능' '병역 관계' '출생지' '본적지' '소속 산업 및 직업'(1937년 7월 1일 중일전쟁 시작 시기와 1940년 10월 1일 현재 기준)을 물었고, 1944년 조사에서도 비슷한 항목이 포함됐다.

6·25 전쟁 후 실시한 1955년 센서스에선 장애 상태를 조사하는 항목이 등장했다. 당시엔 '불구자'란 용어가 쓰였는데, 전쟁 후 부상 정도를 파악하려다 보니 항목도 '실명(失明)-양(兩)·단(單)' '수절단(手切斷)-양(兩)·단(單)' '족절단(足切斷)-양(兩)·단(單)' '농아(聾啞)'로 구성된 식이었다. 1980년에는 심신 장애라는 조사 항목으로 신체뿐 아니라, 정신적 이상으로 생활에 불편을 겪는 가구원이 있는지도 조사되기 시작했다.

문맹 여부를 조사하던 시절도 있었다. 일제강점기인 1930년엔 '가나(仮名·일본어)를 읽고 쓸 수 있나?' '한글(諺文)을 읽고 쓸 수 있나?'를 물었고, 1970년 '글을 읽고

쓸 수 있나?'란 질문을 마지막으로 문맹 여부를 묻는 항목은 사라졌다. 지금은 교육 정도나 전공을 묻고 있다.

영유아 사망이 흔했던 1960년부터는 '총출생아 수'를 묻기 시작했다. 특히 1966년부턴 총출생아 수를 '생존 자녀 수'와 '사망 자녀 수'로 구분해 조사하게 했다. 이는 가장 최근 센서스까지도 포함된 항목이었는데, 올해부턴 '시대착오적'이란 비판 등을 수용해 조사하지 않기로 했다. 2005년에는 저출생 문제가 부각되면서 '추가 계획 자녀 수'도 묻기 시작했다.

집에 '아궁이·변소' 있느냔 질문, 40년 뒤엔 '휴대폰·車'로

인구뿐 아니라 '주택'에 관해서 조사하기 시작한 것은 1960년이다. 당시 부엌·목욕실·광·전등·라디오·외양간 유무를 조사했다. '시멘트' '수세식' '토혈식(土穴式·흙구덩이)' 등 변소 형태나 아궁이 형태도 물었다. 같은 질문도 집의 형태가 바뀌면서 항목이 변화했다. 1970년엔 주택 급수 시설, 1980년엔 목욕 시설 형태를 물었다. 1960년 '개인' '호텔·하숙' '공공건물' 등으로 단출했던 '거처의 종류' 항목도 오늘날엔 '아파트' '오피스텔' 등 10가지로 그 종류가 대폭 늘어났다.

2000년대 들어선 자동차 보유 대수 항목이 들어오기 시작했다. 또 컴퓨터·인터넷 활용 상태와 개인 휴대용 통신기를 사용하는지 여부를 묻는 항목도 등장했다. 개인 휴대용 종류로는 '이동전화기' '무선호출기' '없음'이 제시됐고, 이를 '매일' '일주일에 1번 이상' '한 달에 1번 이상' 등 얼마나 자주 사용하는지도 고르도록 했다.

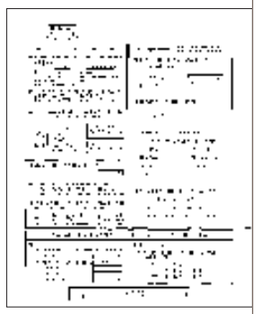
올해 새로운 항목은?… '저출산·돌봄·다문화' 초점

센서스는 기본적인 사회 실태 파악뿐 아니라, 당시의 '주요 정책' 수립을 위해 필요한 자료로서도 수집된다. 이 때문에 통계청은 센서스를 시행하기 전 사회 각계각층과 정부 부처의 요구·자문을 받아 항목 선정에 공을 들인다.

저출산·고령화 이슈가 당면한 큰 위기인 올해엔 '비혼 동거' 항목이 처음 생긴다. 사실혼을 포함해 어떤 혼인 형태도 이루지 않았지만, 함께 살고 있는 가구의 현황을 파악하고자 함이다. 이는 저출생 문제뿐 아니라, 사회의 다양성 문제와도 연관될 수 있다.

통계청 관계자는 "지금은 혼인 신고를 한 대상자 위주로 정책이 만들어지는데, 법적 테두리 밖에 있는 사람을 위한 지원도 필요하다"며 "비혼 동거하는 가구가 어느 정도 규모인지 정확히 측정해 볼 필요가 있다는 공감대가 있었다"고 했다. 이어 "이런 질문을 던짐으로써 '전통적 형태의 가족을 넘어서 확장된 개념의 가족을 받아들여야 할 때가 됐다'는 의식을 국민에게 주고자 하는 의도도 있다"고 덧붙였다.

'가족 돌봄 시간'과 관련한 항목도 구체화해 추가된다. 통계청 관계자는 "장애·건강·고령 이슈로 돌봄이 필요한 사람을 가족 내에서 '누가' '얼마나' 돌보고 있는지 보겠다는 것"이라며 "학교를 가야 하지만 집안에 돌볼 사람이 필요해 학교생활에 집중하지 못하는 '영케어러' 이슈에 대해 파악해 볼 계기가 될 것"이라고 했다. 이밖에 다문화 가정 실태를 파악하기 위해 외국인의 '한국어 실력'도 처음 묻게 된다.



1985 인구 및 주택 센서스 조사표

100년째 기록 완성할 올해 센서스 "꼭 참여해 주세요"

21번째이자 100년째 센서스인 '2025 인구주택총조사'는 올해 10~11월 실시된다. 여기서 도출된 결과는 그간 써 온 역사서에 또 한 장의 보탬이 될 전망이다.

다만 높은 통계의 중요성에 비해 낮은 호응도는 통계청의 고민거리다. 개인 정보를 남에게 선뜻 공개하기 꺼려하는 세태 탓이다. 이 때문에 통계청 사람들은 센서스 질문지의 한 항목을 넣고 빼는 데도 수십 번의 회의를 거친다. 항목 수를 한꺼번에 확 늘릴 수도 없고, 응답자의 심기를 건드리거나 차별적일 수 있는 내용이 들어가서도 안 되기 때문이다.

올해 센서스 항목의 개수는 55개다. 항목에 응답하는 데는 1인당 10분, 2~3인 가구원이라면 20~30분 정도면 충분하다. 통계법 33조에 따르면, 조사되는 내용은 비밀이 엄격히 보장되기에 남에게 누설되지 않으며, 조사 중 알게 된 사항은 통계 생산 목적으로만 사용되고 다른 사람에게는 제공되지 않으므로 안심해도 된다.





지속가능한 사회를 위한 포용적 센서스를 위한 제언

박효민 | 서울시립대학교 도시사회학과 교수

대한민국의 인구주택총조사, 즉 센서스가 이번 2025년을 기점으로 100년을 맞이하게 되었다.

전 세계적으로 국가단위의 인구조사의 역사는 사뭇 오래 되었다고 할 수 있다. 센서스라는 용어가 인구 감찰관(Censor)을 뜻하는 라틴어에서 유래했다는 것에서 나타나듯이, 센서스 제도 자체가 공식적으로 확립된 것은 고대 로마시대이다. 한반도에서도 흔히 우스개 소리로 쓰이는 "호구조사"가 시작된 것은 삼국시대로 거슬러 올라간다. 그리고 대한민국의 근대적 인구주택총조사도 일제강점기 시기인 1925년에 시작되어 이제 100년을 맞이하게 되었다.

애초에 국가에서 영토 내 인구를 자세히 조사하는 것은 정확한 인구 집계를 통해 조세 체계와 병력징집을 효율화하는 것이 가장 큰 목적이 있었다는 것은 잘 알려진 사실이다. 그러나 현대의 민주주의 국가 수립 후 국가의 역할이 크게 변화하였고, 이에 따라 센서스의 목적도 많이 바뀌었다. 현대 민주주의 국가에서는 여러 사회조사를 통해 수집된 자료를 통해 사회상을 파악하고, 이에 대응하는 정확한 정책을 수립하는데 활용하고, 그 중에서도 센서스 자료는 유일한 전수조사라는 의미에서 다른 표본조사로 대체될 수 없는 중요한 역할을 하고 있다.

최근 우리나라를 포함해 각국의 센서스는 크게 두 가지 변화를 겪고 있다.

하나는 센서스 자체의 조사방법의 변화로 기존의 방문면접을 통해 수집하던 자료가 대부분 행정자료로 갈음되고 있다. 이는 점차 대면조사를 기피하는 사회적 분위기의 변화와 함께, 정보 기술의 발전으로 인해 행정자료의 수집, 가공이 용이해 졌기 때문이다. 한국도 지난 2015년부터 일반적 가구특성에 대한 정보는 등록센서스를 이용해 수집하고 있다. 이는 국가의 조사비용과 더불어 응답자의 조사 부담을 줄이는데 큰 기여를 하고 있다.

다른 한 가지는 센서스의 대상이 되는 사회적 환경의 변화이다. 사회가 개인화되고 지역 공동체의 결속이 약화하며, 동시에 다양한 가치관과 배경을 지닌 사람들이 공존하는 사회로 변화해 가면서 센서스의 성격 역시 변화를 요구하고 있다. 사회적 환경이 변화함에 따라 사회조사에서 이전에는 사회적 관심이 되지 않았던 문항을 수집할 필요가 생기거나, 혹은 기존의 문항에서도 세분화되고 다양한 데이터를 수집할 필요가 높아지고 있다.

센서스가 이러한 변화에 부응하는 것은 사회의 지속가능성과 포용성을 높이는데 중요하다. 따라서 우리도 다양한 사회집단에 대한 통계자료를 생성함으로써 보다 포용적인 센서스의 방향을 모색할 시점이 된 것이다.

센서스가 당면한 가장 중요한 사회문제 중 하나는 인구감소 문제, 그리고 맞물려 이주배경 주민이 증가하고 있다는 점이다.

특히 최근들어 비교적 단기간 안에 다양한 이주배경 주민이 증가하고 있으며, 이들과 원주민의 상생과 통합이 사회적 당면문제로 떠오르고 있다. 따라서, 이를 위한 기초자료 생성이 그 어느 때 보다 중요한 시점이다. 이와 같은 배경 하에 센서스에서도 사회구성원들의 이주배경에 대한 문항이 앞으로 더욱 늘어나고 세분화 되어야 필요가 있다.

한국의 센서스에는 이주관련 문항이 최소한으로 포함되어 있었는데, 이는 한국이 그 동안 경험해 온 인종과 민족의 단일성 담론에 기인한 것으로 보인다. 그 결과 한국의 센서스는 이주배경에 대해 응답자의 국적과 출생지, 입국년도 정도만을 묻고 있다.

이는 1990년대 이후 본격적으로 유입되기 시작한 이주노동자, 재외한인동포 등에 대한 자료를 생성할 수 있을 정도의 질문이다. 그러나 이제 대한민국의 이주의 역사가 길어지고, 수가 증가하며, 동시에 이주 과정의 다양화가 일어나면서, 이주 2세나 3세 등이 늘어나고 있으며, 또한 겉보기로 나타나는 인종적 차이와 자신의 문화적 정체성이 다르게 형성되는 경우가 늘어나고 있다.

따라서 이주배경 문항을 확대하여 이 변화를 보다 정밀하게 파악하려는 노력이 필요하다.

외국 센서스의 경우 우리와 이주에 대한 배경과 역사, 경험이 다르기 때문에 그대로 비교하여 참고할 수는 없으나, 이주배경 문항에 대해 보다 자세하게 질문하는 경우가 많다. 예를 들어 호주 센서스의 개인용 설문에서는 응답자 자신의 출생지 뿐만 아니라, 부모의 출생지를 질문하고 있으며, 이에 더해 본인 스스로를 어떤 국가 혹은 문화권에서 유래한 사람이라고 생각하는지를 응답하도록 하고 있다. 영국의 센서스 역시 스스로 생각하는 자신의 민족적 배경을 매우 다양한 각도로 묻고 있다. 또한 영국, 캐나다 등의 센서스에서는 주로 사용하는 언어와 사용할 수 있는 언어 등을 질문하여 사회 구성원들의 언어 사용도 확인하고 있다. 물론 이는 호주나 영국, 캐나다 등의 국가들이 식민지 체제를 겪은 역사와 관련이 있는 것으로 보이며, 우리와는 매우 다른 상황이다. 그럼에도 불구하고 이와 같은 이주문항의 다양성은 다문화 국가로 향하고 있는 한국 역시 이주민 여부나 출신국의 정보를 넘어서 스스로 정체성을 가지는 민족적, 문화적 배경을 조사할 필요가 있다는 점을 시사한다.

종교적 소수자에 대한 세분화된 자료 생산도 가능하다.

현재 한국의 센서스에서는 주요 종교를 중심으로 질문을 하고 있으며, 이는 지난 수 십년간 주로 흔히 말하는 3대 종교를 중심으로 한국사회의 종교가 구성되어 있었기 때문에 이를 반영한 것으로 볼 수 있다. 그러나 최근 종교 내에서 다양한 성격으로 분파가 갈라지고 있으며, 특히 대한민국에서 두드러지게 나타나지 않았던 종교인들의 숫자도 크게 늘어나고 있는 상황이다.

2015년 센서스에서는 기타 종교를 가진 사람들을 98,185명으로 집계하고 있으나, 많은 연구에서 기타 종교의 범주에 증가세가 확인되고 있다. 예를 들어 이슬람교의 경우 한국인 신자가 이미 6만명 수준으로 제시되고 있으며, 장, 단기 체류 외국인을 포함하면 이슬람 신자는 수십 만명 수준으로 추산되고 있다. 그러나 이에 대해 추정할 수 있는 국가 통계가 없기 때문에, 여러 다른 수치들이 제시되고 있는 상황이다. 현행 센서스는 10년 주기로 종교를 파악하고 있는데, 이때 기타 항목을 보다 세분화 하여 정확한 종교 분포를 파악한다면 한국사회의 종교적 분포에 대해 보다 정확한 자료를 생성할 수 있을 것이다. 또한 현재에는 큰 의미가 없는 숫자일지라도, 앞으로 변화 추세를 살펴보기 위해서는 지금부터 자료를 수집, 보관할 필요가 있다.

또 다른 방향에서 보다 세분화된 자료를 생성할 수 있는 것은 혼인 및 출산 관련 문항이다.

다른 표본조사에서도 가족구성이나 출산관련 자료를 모으고 있으나, 20% 표본조사인 센서스가 혼인, 출산 패턴을 면밀하게 파악하여야 전 국가적으로 관심을 기울이고 있는 저출산 문제에 보다 효과적으로 대응할 수 있을 것이다. 우선 현재 한국 센서스에서는 출산 여부를 질문하고 있으나, 미혼으로 응답한 응답자에게는 출산 여부를 묻지 않게 스킵패턴(skip-pattern)을 적용해 통해 출산 문항을 넘어가도록 설계가 되어 있다는 점이 눈에 띈다. 이는 전통적 가족구조를 염두한 질문 구성이지만, 한국사회의 전통적 가구구성이나 출산 형태가 크게 변화하고 있는 만큼 혼인 여부와 출산여부는 두 문항을 연관 짓지 않고, 독립적으로 구분해서 질문할 수 있다.

비혼출산은 앞으로 한국사회의 인구구조에 큰 영향을 미칠 수 있다. 한국은 비혼출산에 대한 공적 지원이 발달하지 않았으며, 사회적 시선 역시 부정적인 경우가 있다. 그 결과 2022년 기준 혼외출산율이 2.9%인 것으로 집계되고 있다.

그러나 비혼출산에 대한 사회적 지원이 제공되고 사회적 인식 역시 부정적이지 않은 유럽연합의 몇몇 국가들은 비혼출산의 비율이 많게는 절반 이상을 기록하고 있으며, 이들 나라는 한국에 비해 월등하게 높은 출산율을 보이고 있다. 이러한 점에서 비혼출산의 파악은 저출산 정책 방향에 큰 영향을 미칠 수 있다.

이와 더불어 혼인관계에 대한 파악도 보다 다각도로 이루어 져야 할 것이다. 2020년에 시행된 한국의 센서스는 법적 혼인관계에 있으나 사실상 별거로 인해 혼인관계가 해소된 상태 (직장 등으로 인해 따로 사는 것은 제외)나, 반대로 실제 법적 혼인관계는 아니지만 사실혼 관계에 있는 사람들을 파악하지 못한다는 한계가 있다. 특히 2015년 센서스 문항에서는 혼인상태를 ① 미혼, ② 배우자 있음, ③ 동거, ④ 사별, ⑤ 이혼, ⑥ 별거 등 6개의 범주로 응답하도록 하였으나, 2020 센서스에서는 혼인상태에 대한 응답지가 ① 미혼, ② 배우자 있음, ③ 사별, ④ 이혼 등 4개의 범주로 축소되어 세분화된 자료를 얻기 어렵게 되었다.

반면, 많은 국가들에서 혼인상태는 매우 자세하게 파악하려 노력하고 있다. 예를 들어 영국의 경우에는 9개의 범주로 물어보고 있으며 특히 그 중에는 이성 관계와 동성 관계 모두에게 허용되는 생활동반자관계(civil partnership) 여부나 배우자의 동성 혹은 이성 여부도 자세하게 물어보고 있다. 한국의 경우 아직 생활동반자 관계에 대해 법적으로 지위가 부여되고 있지 않으며, 동성간 결혼도 법적으로 허용되지 않고 있어 이와 같은 설문을 센서스에 포함시키는건 현재로서는 어려움이 있지만, 전반적으로 혼인상태에 관련된 문항에 대한 다양화는 고려해 볼 수 있을 것이다. 외국의 센서스는 혼인과 성 정체성에 대해 자세한 문항 뿐만 아니라 다양한 선택지를 주는 경우도 있다. 캐나다의 경우 출생시 등록된 성(sex)과 현재 사회적으로 나타내는 성(gender)을 별도의 문항으로 따로 질문하고 있다. 영국 센서스의 경우 이성애, 동성애, 양성애 등의 선택지를 통해 성적 지향을 질문하고 있으며, 이와는 별도로 생물학적 성과 사회적 젠더가 일치하는지 여부를 질문하고 있다. 그러나 이는 응답과정에서의 민감한 개인정보를 질문한다는 점에서 윤리적 문제 혹은 조사의 어려움을 야기할 수 있다. 영국의 경우 이와 같은 성적 지향에 대한 질문을 할 경우 민감한 질문들은 대답하고 싶지 않은 경우 응답을 이끌어 내지 않고 응답자의 자율에 맡기는 '자발적 응답 질문(voluntary question)'으로 구분하여 응답을 받고 있다.



마지막으로 직업의 여러 형태에 대해 질문 할 수 있는 방안들도 고려할 수 있을 것이다.

정보통신 기술의 발전과 직업 구조의 변화로 인해 기존의 전통적 노동과는 다른 방식의 노동형태 혹은 노동계약 형태가 증가하고 있다. 예를 들어 플랫폼 노동의 증가는 한국사회의 큰 변화로 나타나고 있으며, 고용노동부에 따르면 2023년 플랫폼종사자의 규모는 전국적으로 88.3만명으로 추산되고 있다. 이는 그 전해에 비해 11.1%가 증가한 것이고, 3년전에 비해 30% 이상 증가한 것으로 매우 가파른 증가세를 보이고 있다. 이러한 노동구조의 변화는 한국사회 전반에 걸쳐 매우 큰 영향을 미칠 수 있기 때문에, 플랫폼 노동, 비정규적 고용 등의 실태에 대해 센서스를 통해 파악한다면, 노동시장 정책 수립에 도움이 될 수 있을 것이다.

센서스와 같은 정기적 조사는 동일한 문항을 통해 시계열적 안정성을 확보하는 것이 무엇보다 중요하다.

따라서 센서스의 문항을 추가하거나, 응답지를 변경하는 것은 매우 신중을 기해야 하는 일이다. 그러나 다른 한편으로 커다란 사회적 변화를 겪고 있는 상황에서 센서스의 효용성을 높이기 위해서는 장기적 차원에서 센서스의 틀 변화도 필요하다. 특히 사회통합과 지속가능성의 확보를 위한 센서스 문항의 확대 및 세분화가 앞으로 센서스의 또 다른 100년을 준비하는 시점에서 이루어 지길 희망한다.

ISSUE 04

인구주택총조사 100년: 센서스의 의의와 향후 발전방안

전광희 | 충남대학교 사회학과 명예교수



인구주택총조사는 한 나라의 인구와 주택에 관한 가장 기본적이고 포괄적인 통계를 수집하는 국가적 사업이다. 유엔 통계처(UNSD)는 인구센서스를 "특정 시점에 한 국가 또는 명확히 구분된 지역 내의 모든 사람에 관한 인구, 경제, 사회 자료를 수집, 집계, 분석 및 배포하는 전체 과정"으로 정의한다.

한국의 경우, 1925년 일제 강점기에 첫 인구센서스가 실시된 이후, 해방 후 1949년에 대한민국 정부 수립 후 첫 자주적 센서스를 시행했으며, 1960년부터 유엔의 글로벌 기준에 맞춰 5년마다 정기적으로 실시해오다가, 2015년부터 행정자료를 이용하는 등록센서스와 20% 표본조사를 결합하는 혼합센서스로 전환하였다.

인구주택총조사의 중요성과 필요성

통계학적 관점에서의 중요성

인구주택총조사는 전수조사로서 표본오차가 발생하지 않는다는 장점이 있어 소지역 통계나 소집단 통계 작성에 있어 중요한 의미를 지닌다. 또한 표본조사의 추출 프레임을 제공함으로써 국가통계 체계의 근간으로 기능한다.

정책적 관점에서의 필요성

증거기반 정책 수립의 토대로서, 인구의 규모, 분포, 특성에 관한 정확한 정보는 교육, 보건, 주택, 교통, 복지 등 다양한 공공정책 영역에서 자원 배분과 서비스 설계의 기준이 된다. 특히 지방분권화가 진전됨에 따라 지역별 맞춤형 정책 수립에 인구주택총조사 자료의 중요성이 더욱 커지고 있다.

사회경제적 관점에서의 가치

인구주택총조사는 저출산, 고령화, 가구구조 변화, 도시화, 지방소멸 등의 현상을 시계열적으로 추적할 수 있게 함으로써 사회변동 연구의 실증적 기반을 제공한다. 또한 민간 기업의 경영전략 수립과 노동시장, 금융시장, 부동산 시장의 기초 분석에도 활용한다.

인구주택총조사의 활용 사례

정부 정책에서의 활용	**중앙정부**는 국토종합계획, 저출산·고령사회 기본계획, 주택종합계획 등의 수립에 인구주택총조사 자료를 활용한다. 지방교부세 산정의 기준으로도 활용되어 국가 자원 배분의 형평성을 확보하는 역할을 한다. **지방자치단체**는 도시계획, 주택정책, 교통계획 등 지역 발전 전략의 근간으로 인구분포와 특성 자료를 활용한다. 인구소멸 위기 지역 선정, 청년 인구 유입 정책, 노후주택 개선사업 등 지역 맞춤형 정책 개발에 활용된다.
민간 부문에서의 활용	**유통업체**들은 인구밀도, 연령분포, 가구특성 등을 고려하여 매장 입지와 상품구성을 결정한다. **금융기관**은 지역별 인구구조와 소득수준을 분석하여 맞춤형 금융상품을 개발한다. **부동산 개발업체**들은 가구형태와 연령대별 인구전망을 바탕으로 주택수요를 예측하고 개발계획을 수립한다.
공공서비스 설계에서의 활용	**교육청**은 학령인구 변화를 예측하여 학교 신설·통폐합 계획을 수립하고, **보건소**는 지역 인구구조에 맞는 건강증진 프로그램을 개발한다. **소방서**와 **경찰서**는 인구밀도, 주택유형, 연령구조 자료를 활용하여 재난대응 계획을 수립한다.
재난 및 위기 대응에서의 활용	**코로나19 팬데믹 상황**에서 보건당국은 인구밀도, 연령구조, 가구 특성 자료를 활용하여 감염취약지역을 예측하고 선제적 방역 전략을 수립했다. **자연재해 대응과 기후변화 대응 정책**에서도 인구주택총조사 자료를 중요하게 활용한다.

EUROPE



북유럽 국가들은 센서스 방법론 혁신의 선두주자로, 덴마크(1981년), 핀란드(1990년), 스웨덴(1995년), 노르웨이(2011년) 등이 순차적으로 등록센서스로 전환했다. 이들 국가는 발달된 주민등록시스템, 부동산등록, 교육·고용 데이터베이스 등을 연계하여 현장조사 없이도 정확한 센서스 통계를 생산하고 있다. 특히 덴마크는 40년 이상의 등록센서스 경험을 바탕으로 다른 국가들에게 기술적 조언을 제공하고 있다.



서유럽에서는 네덜란드가 2011년부터 행정자료와 표본조사를 결합한 가상 센서스(virtual census) 방식을 도입했다. 독일은 2011년에 주거지등록부와 표본조사를 결합한 방식으로 센서스를 실시했으며, 2022년 센서스에서는 행정자료의 활용 범위를 더욱 확대하고 있다. 프랑스는 2004년부터 순환식 센서스(rolling census) 방식을 도입하여 매년 20%의 지역을 조사하고 5년 주기로 전체 결과를 집계하는 방식을 채택하고 있다.



남유럽에서는 이탈리아가 2018년부터 영구 센서스(permanent census) 방식을 도입했다. 이는 매년 표본조사를 실시하고 행정자료와 결합하여 연간 인구통계를 생산하는 방식으로, 전통적인 10년 주기의 센서스에 비해 시의성이 크게 향상되었다. 스페인은 2021년 센서스에서 주민등록부와 행정자료를 기반으로 한 등록센서스로 전환했으며, 그리스는 2021년 센서스에서 온라인 조사를 처음으로 도입하여 디지털 전환을 추진했다.



ASIA



일본은 1920년부터 5년마다 정기적으로 센서스를 실시하고 있으며, 2020년에는 코로나19 상황에서도 97.7%의 높은 회수율을 기록했다. 일본은 인구감소와 고령화라는 인구학적 위기 속에서 센서스 자료를 지방창생(地方創生) 정책의 핵심 근거자료로 활용하고 있다.



싱가포르는 2020년에 완전한 등록센서스로 전환했다. 국민등록, 주택, 교육, 고용 등의 행정자료를 연계하여 센서스 정보를 생산하고, 표본조사를 통해 행정자료에서 얻기 어려운 정보를 보완하는 방식을 채택했다. 이는 동아시아 국가 중 행정자료 기반 센서스로의 전환을 선도하는 사례로 평가된다.



대만은 2020년 센서스에서 인터넷 조사와 행정자료를 결합한 하이브리드 방식을 도입했다. 대만 통계청은 행정자료의 활용도를 높이면서도 전통적인 현장조사의 장점을 유지하는 절충적 접근을 취하고 있다. 특히 가구구성과 주거환경에 관한 상세 정보를 수집하기 위해 표본조사를 병행하는 방식이 주목받고 있다.



태국은 2020년 센서스에서 태블릿 기반의 전자조사를 도입하여 자료수집 과정의 디지털화를 추진했다. 태국 통계청은 센서스를 통해 수집된 자료를 국가 개발계획과 지역 균형발전 정책에 적극 활용하고 있으며, 특히 도시-농촌 간 발전 격차 해소를 위한 기초자료로 센서스 결과를 활용하고 있다.



이스라엘은 2008년부터 등록센서스 방식을 도입했으며, 2020년 센서스에서는 행정자료와 표본조사를 결합한 통합적 접근법을 채택했다. 이스라엘 통계청은 인구등록부, 주소등록부, 건축물대장 등 다양한 행정자료를 통합하여 기본 인구통계를 생산하고, 이를 보완하기 위한 대규모 표본조사를 실시하는 방식으로 센서스를 수행하고 있다.



터키는 2021년 센서스에서 주소기반 인구등록시스템을 활용한 등록센서스로 전환했다. 터키 통계청은 주소기반 인구등록시스템을 중심으로 다양한 행정자료를 연계하여 센서스 정보를 생산하고 있으며, 특히 지역개발 정책과 도시계획 수립에 센서스 자료를 적극 활용하고 있다.

학계에서의 인구주택총조사 활용

인구주택총조사는 학술 연구의 핵심 자료로서 다양한 학문 분야에서 활용되고 있다. 특히 인구주택총조사의 전수조사 특성과 시계열적 연속성은 한국 사회의 구조적 변화를 분석하는 귀중한 자료로 인정받고 있다.

인구학 및 사회학 연구

인구학 분야에서는 인구주택총조사 자료를 기반으로 한 인구추계와 인구구조 변화 연구가 활발히 이루어지고 있다.
또 한국의 출산율 변동 요인을 코호트 관점에서 분석하는 연구들이 진행되고 있으며, 이를 통해 세대별 출산 행태의 차이와 지역 간 출산율 격차의 구조적 원인을 규명하고 있다.
사회학 분야에서도 가족구조 변화에 관한 연구에 인구주택총조사 자료를 활용한다.
1인 가구의 증가가 지역사회에 미치는 영향을 분석하고, 특히 고령 1인 가구와 청년 1인 가구의 공간적 분포와 사회경제적 특성을 비교함으로써, 세대별로 차별화된 사회서비스의 필요성을 강조하는 연구가 수행되고 있다.



경제학 및 불평등 연구

경제학 분야에서는 소득 및 자산 불평등 연구에 인구주택총조사 자료를 활용한다.
인구주택총조사 자료와 가계금융복지조사를 연계하여 세대별, 지역별 자산 불평등 구조를 분석하는 연구가 진행되고 있다.
특히 주택 소유 여부에 따른 자산 격차가 2000년대 이후 급격히 확대되는 현상을 밝히고, 이것이 세대 간 자산 이전과 결합하여 불평등의 상속으로 이어지는 메커니즘을 규명하는 연구가 이루어지고 있다.
교육 분야에서는 교육기회의 불평등 연구에 인구주택총조사 자료를 활용한다.
인구주택총조사 자료의 시계열적 변화를 분석하여 한국 사회의 교육 수준 향상이 사회이동성과 소득 불평등에 미친 영향을 연구한 사례가 있다.
이러한 연구는 교육 기회의 확대가 세대 간 사회이동에 미친 긍정적 효과와 한계를 실증적으로 분석하고 있다.



도시계획 및 지리학 연구

도시계획 분야에서는 공간구조 분석에 인구주택총조사 자료가 활용된다.
인구주택총조사의 통근·통학 자료를 활용하여 대도시권 공간구조 변화를 분석한 연구들이 진행되고 있다.
이러한 연구는 수도권 직주분리 현상의 심화와 교통망 확충에 따른 생활권 변화를 실증적으로 보여주고 있다.
지리학 분야에서는 인구 현상의 공간적 분포 연구에 인구주택총조사 자료가 활용된다.
인구주택총조사 자료를 GIS와 결합하여 인구 고령화의 공간적 전개 과정을 시각화하는 연구가 수행되고 있다.
이러한 연구는 농산어촌 고령화에서 시작하여 소도시, 중소도시로 확산되는 고령화의 공간적 패턴을 밝혀내고 있다.



인구주택총조사 참여의 중요성

정확한 통계 생산의 기반

올해 2025년 인구주택총조사에서, 특히 20% 표본조사의 정확성과 신뢰성은 국민들의 적극적인 참여에 달려있다. 비응답이 증가할 경우, 20% 대규모 표본조사의 결과도 통계의 대표성이 저하되고, 이는 정책 수립의 왜곡으로 이어질 수 있다.

사회적 포용성 제고

인구주택총조사는 소외된 집단을 포함한 모든 국민의 현실을 반영하는 유일한 통계조사로, 모든 국민의 참여는 포용적 사회 발전을 위한 기초가 된다.



지역사회 발전의 토대

인구주택총조사 자료는 지방자치단체의 재정 배분과 공공서비스 공급 계획의 기준이 되며, 정확한 인구통계는 지역맞춤형 지속가능 발전 전략 수립에 필수적이다.



인구주택총조사의 미래와 과제

디지털 전환과 방법론적 혁신

온라인 조사와 모바일 조사의 확대, 행정자료 활용 강화, 빅데이터 연계 등이 주요 변화 방향이다.
한국의 인구주택총조사도 교육, 고용, 보건 등 다양한 분야의 행정자료 활용을 점진적으로 확대하고, 궁극적으로는 등록센서스로의 완전한 전환을 목표로 할 필요가 있다.

데이터 활용성 강화와 개방성 확대

데이터 시각화, API 제공, 마이크로데이터 접근성 향상 등을 통해 인구주택총조사 자료의 활용 가치를 극대화할 필요가 있다.
동시에 개인정보 보호와 데이터 개방 사이의 균형도 중요한 과제이다.



초저출산, 고령화, 인구감소 시대의 총체적 통계적 대응

향후의 인구주택총조사에서는 지방소멸, 인구 고령화, 가구구조 변화 등 인구감소에 수반되는 다양한 사회경제적 변동의 추이를 시의적절하게 측정하고, 이에 총체적으로 대응하는 맞춤형 정책 수립의 근거를 확실하게 제공하는 것이 대단히 중요한 과제이다.

결 론

인구주택총조사 100년의 역사는 한국 사회의 급격한 변화와 발전을 객관적 수치로 기록해온 소중한 여정이다. 앞으로 인구주택총조사가 직면한 과제는 디지털 전환의 가속화, 행정자료의 효과적 활용, 인구감소 시대의 통계적 대응 등 다양하다.
인구통계는 궁극적으로 국민의 삶의 질 향상을 위한 정책 수립의 기초자료로서 가치를 갖는다. 인구주택총조사가 앞으로도 한국 사회의 현재를 정확히 진단하고 미래를 준비하는 귀중한 자산으로 기능하기 위해서는, 국민 모두의 관심과 참여가 필수적이다. 또한 정부와 학계, 민간 부문의 협력을 통해 인구주택총조사 자료의 활용 가치를 극대화하고, 한국 사회의 지속가능한 발전을 위한 통계적 기반을 더욱 강화해 나가야 할 것이다.

AI가 바꾸는 교실: 새로운 리터러시의 시대

김용성 | 충남대학교 기술교육과 교수



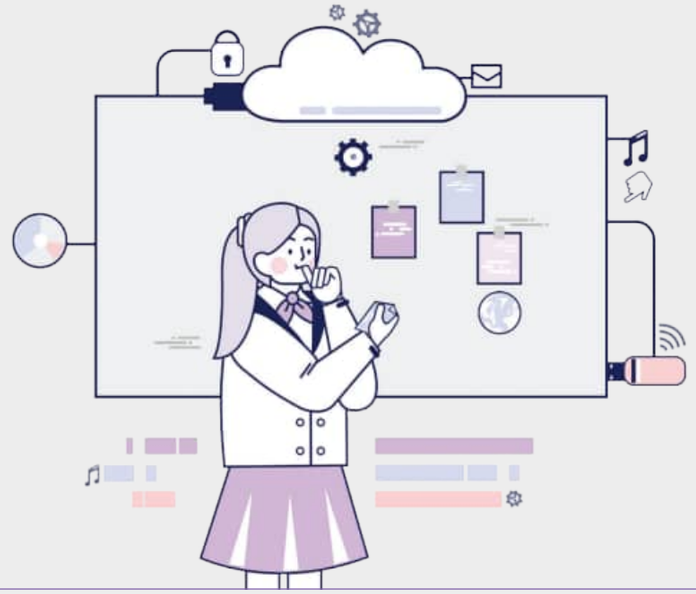
생성형 AI와 교육의 미래

학교 현장에서 챗GPT가 없으면 가장 불편함을 느끼는 사람은 누구일까?
최근 교사 대상 연수에서 필자는 흥미로운 질문을 던진 적이 있다.
"혹시 이번 생활기록부 작성하실 때, 생성형 AI를 전혀 활용하지 않으신 분 계신가요?" 놀랍게도, 약 80%의 교사들이 생성형 AI를 '적극적으로' 혹은 '부분적으로' 활용했다고 답했다. 물론 해당 연수장은 AI 활용 경험이 많은 교사들이 모인 자리였기에 응답 비율이 다소 높았을 수는 있다.
하지만 단일 학교가 아닌 여러 학교에서 참여한 교사들이 공통적으로 생성형 AI의 편리함을 실감하고 있다는 점에서, 교육 현장에서의 생성형 AI 확산은 더 이상 예외적 현상이 아니다.

그렇다면 이런 편리한 도구가 사라진다면 누가 가장 불편을 느낄까? 위 질문에 따르면 가장 불편함을 느끼는 사람들은 단연 선생님으로 생각할 수 있다. 하지만, 더 불편함을 느낄 사람들은 학생들일 수 있다. 실제로 한국청소년정책연구원(2024)의 전국 청소년 대상 조사 결과, 청소년의 52.1%가 생성형 AI를 이용한 경험이 있다고 답했으며, 이들은 주로 정보 검색(3.93점)과 과제 수행(3.59점)에 AI를 활용한다고 밝혔다. 사용 이유는 '관심과 호기심'이 59.5%로 가장 높았고, 하루 평균 이용 시간은 '30분 미만'이 75.7%로 집계되었다. 이는 생성형 AI가 청소년들의 일상에도 빠르게 자리잡고 있다는 것이다. 현재 학생들은 단순한 과제 수행부터 복잡한 정보 검색까지 생성형 AI의 도움을 받으며 의존도가 점점 높아지고 있다. 요즘엔 단순한 구글 키워드 검색을 해도 구글의 Gemini의 답변도 함께 보여주기 때문에 학생들은 자연스럽게 AI에 노출이 될 수 밖에 없다. 이제는 피할 수 없는 AI, 학교 현장에서 어떻게 하면 현명하게 활용할 수 있을까? 그리고 활용하면서 큰 문제는 없을까? 이번 원고에서는 다양한 생성형 AI로 인해 우리 교육 현장에서 생각해보아야 할 문제에 대해서 서술해보고자 한다.

AI 리터러시의 중요성

디지털 리터러시를 넘어 AI 리터러시가 새로운 필수 역량으로 부상하고 있다. AI 리터러시란 단순히 AI를 사용하는 기술적 능력을 넘어, AI의 작동 원리를 이해하고, 그 결과물을 비판적으로 평가하며, 윤리적으로 활용할 수 있는 종합적 역량을 의미한다. 이는 AI 시대를 살아갈 학생들에게 더 이상 선택이 아닌 필수가 되었다.
AI를 효과적으로 활용할 수 있는 역량은 미래 사회에서 핵심 경쟁력이 될 것이다. 복잡한 문제 해결, 방대한 정보 처리, 창의적 아이디어 발상 등 다양한 영역에서 AI는 인간의 능력을 확장시키는 강력한 도구가 될 수 있다. 특히 교육 분야에서는 학생 개개인에게 맞춤형 학습을 제공할 수 있고, 학생들의 학습 과정을 분석하여 효율적인 교육 방법을 제공하는 데 큰 도움이 될 수 있다. 따라서 학생들이 자신의 학습 목표에 맞게 적절한 AI 도구를 선택하고, 이를 창의적으로 활용할 수 있는 능력을 기르는 것은 무엇보다 중요한 일이다.
그러나 학생들에게 무조건적인 AI 활용을 권장하는 것은 상당히 위험할 수 있다. 대부분의 생성형 AI 서비스들은 연령 제한을 통해 미성년자들의 AI 사용을 제한하고 있다. 따라서 미성년자 학생들에게는 AI 서비스 사용을 적절히 조절할 수 있도록 교육이 필요하다. 무분별하게 AI 활용법만 가르치다 보면, 가장 창의력이 풍부한 시기에 학생들이 스스로 사고하고 문제를 해결



하는 기회가 제한되어 결과적으로 사고력 발달에 악영향을 미칠 수 있다.
또한, AI가 제공하는 정보를 무비판적으로 수용하는 것의 위험성을 인식시키는 교육이 필요하다. 생성형 AI는 때로 사실과 다른 정보(환각)를 생성하거나, 특정 데이터에 편향된 결과를 도출할 수 있다. 학생들은 AI가 제시한 정보의 출처를 확인하고, 다른 자료와 교차 검증하는 습관을 길러야 한다. 이는 단순한 기술 교육이 아닌, 비판적 사고력과 정보 판별 능력을 기르는 근본적인 교육 과정이 되어야 한다.
AI를 활용한 결과물에 대한 윤리적 책임 의식도 중요하다. 학생들은 AI를 통해 얻은 결과물을 어떻게 인용하고 활용해야 하는지, 또 어디까지가 허용 가능한 활용 범위인지에 대한 명확한 가이드라인을 배워야 할 필요가 있다.
교사들 역시 AI 리터러시 함양이 시급하다. 교사가 AI에 대한 이해도가 부족하면, 학생들의 AI 활용을 적절히 지도하거나 평가할 수 없기 때문이다. 교육청과 학교는 교사들을 위한 체계적인 AI 연수 프로그램을 확대하고, 현장에서의 실제 적용 사례(수업, 업무 활용 사례 등)를 공유할 수 있는 체계를 마련해야 한다. 또한, 교사들이 AI 도구를 직접 체험하고 실무에 활용할 수 있도록 AI 서비스 사용료 지원 등의 정책적 지원을 마련하는 것도 매우 중요하다.
결국 AI 리터러시 교육의 목표는 학생들이 AI를 단순한 도구로 활용하는 것을 넘어, AI와 함께 성장하고 공존하는 방법을 배우게 하는 것이다. 이는 미래 사회에서 학생들이 AI를 주체적으로 활용하며 자신의 역량을 확장할 수 있는 근본적인 토대가 될 것이다.

생성형 AI 시대의 평가 패러다임의 전환

생성형 AI가 교육 환경에 깊숙이 들어오면서 전통적인 학생 평가 방식에 근본적인 변화가 필요해졌다.
"이 과제가 정말 학생 스스로만의 노력으로 얻은 결과인가?"라는 질문 자체가 무의미해지는 시대가 되었다. 거의 모든 학습 과정과 결과물에 AI의 영향이 어떤 형태로든 개입될 수 있기 때문이다. AI의 도움 없이 순수하게 학생이 만든 산출물을 구분하는 것은 갈수록 어려워지고 있으며, 이는 수십 년간 우리 교육이 지켜온 평가 체계의 근본적인 변화가 불가피함을 보여준다.
기존의 과제 평가 방식은 주로 최종 결과물에 초점을 맞추었다. 결과물의 완성도, 내용의 충실도 등이 주요 평가 대상이었다. 그러나 이제 이러한 결과물은 생성형 AI의 도움으로 비교적 쉽게 높은 수준으로 작성될 수 있다.
이런 상황에서 과거의 평가 방식을 고수한다면, 실제로 평가하는 것은 학생의 본질적 능력이 아닌 AI 활용 능력이 되어 주객전도 상황이 되어버리는 것이다.
따라서 새로운 평가 패러다임은 '과정'과 '활용'에 초점을 맞추어야 한다. 예를 들어, AI에게 어떤 질문을 했는지, 그 결과를 어떻게 편집하고 활용했는지, 여러 출처에서 얻은 정보를 어떻게 통합했는지 등의 과정을 평가하는 것이 더 의미 있는 접근이 될 수 있다. 즉, 학생이 AI를 얼마나 창의적이고 효과적으로 활용했는지가 평가의 핵심이 될 수 있는 것이다.
실시간 평가의 중요성도 커질 것이다. 학생들이 직접 교실에서 문제를 해결하는 과정을 관찰하거나, 실시간으로 사고 과정을 설명하게 하는 방식의 평가가 더욱 높은 가치를 갖게 될 것이다. 이는 단순한 지식 암기나 정보 수집을 넘어, 비판적 사고력, 문제 해결 능력, 의사소통 능력 등 AI가 쉽게 대체할 수 없는 고차원적 사고 능력에 초점을 맞추는 평가 방식이다.
또한, AI를 적극적으로 활용하는 평가 방식도 고려할 수 있다. 예를 들어, "AI가 한 답변을 비판적으로 분석하고, 새로운 형태로 글을 써보세요"와 같은 과제는 학생들의 비판적 사고력을 평가하는 동시에, AI를 교육적 도구로 활용하는 방법을 가르칠 수 있다.
결국 생성형 AI 시대의 교육 평가는 '무엇을 알고 있는가'보다 '무엇을 할 수 있는가'에 더 큰 가치를 부여해야 한다. 이는 지식 기반 사회에서 역량 기반 사회로의 전환을 반영하는 것으로, 교육의 본질적 목표에 더 부합하는 방향이라고 할 수 있다.



AI와 함께하는 미래 교실의 모습

생성형 AI가 교육 현장에 본격적으로 도입된다면, 교실은 어떻게 변화할까? 올해부터 초중등 교육 현장에서는 이미 AI 디지털교과서(AIDT)가 일부 도입되어 학생들이 활용하고 있다. 이제 교실에서의 AI 활용은 더 이상 먼 미래의 이야기가 아니다. 이는 단순히 새로운 기술 도구의 도입이 아닌, 교육 방법론과 교실 문화의 근본적인 변화를 의미한다. AI와 공존하는 교실 모델은 교사와 학생 모두에게 새로운 가능성을 제시할 수 있을 것이다.

가장 먼저 학교 현장에서는 교과별로 AI 활용 전략을 수립하고, 각 과목의 특성에 맞게 차별화된 접근법을 개발해야 할 것이다. 예를 들어, 국어 수업에서는 AI를 활용한 글쓰기 피드백 시스템이나 텍스트 분석 도구를 활용할 수 있다. 수학에서는 학생이 직접 문제를 풀어보고, AI를 활용하여 다른 방식으로 문제를 풀 수 있는 방법에 대해 함께 논의해볼 수 있다. 사회 과목에서는 역사적 사건을 다양한 관점에서 해석하는 데 AI를 활용할 수도 있다. 이처럼 앞으로 AI는 각 교과의 본질적 학습 목표를 더 효과적으로 달성할 수 있는 방향으로 활용되어야 할 것이다.

AI를 활용한 개인 맞춤형 학습은 오래전부터 이상적인 교육 모델로 여겨져 왔다. 그동안 공교육에서는 이러한 맞춤형 학습을 실현하기 어려웠지만, 새로운 형태의 AI 기반 교과서가 등장하면서 '모든 학생에게 최적화된 교육'을 구현할 가능성이 높아지고 있다. 학생별 학습 속도, 이해도, 관심사에 따라 AI가 맞춤형 학습 자료와 피드백을 제공하면, 교사는 더 많은 시간을 학생과의 직접적인 상호작용이나 정서적 지원에 할애할 수 있다.

교실 운영 방식도 변화할 것이다. 전통적인 일방향적 강의 중심 수업은 감소하고, 협력적 문제 해결이나 프로젝트 기반 학습이 증가할 것이다. 예를 들어, 학생들은 AI의 도움을 받아 필요한 정보를 빠르게 수집하고 정리한 후, 이를 바탕으로 더 깊이 있는 토론이나 창의적 활동에 시간을 투자할 수 있다. 이런 과정에서 교사는 정보의 정확성을 검증하고, 다양한 관점을 제시하며, 학생들의 비판적 사고를 자극하는 조력자 역할을 수행하게 될 것이다.

그러나 이러한 변화는 교사들에게 새로운 역량을 요구한다. 기존의 교수법을 넘어, AI 도구의 적절한 선택과 활용, 학생들이 AI를 효과적으로 활용할 수 있도록 하는 지도 방법, 그리고 새로운 형태의 수업 설계 능력 등이 필요하다. 이는 교사 양성 과정과 연수 프로그램의 혁신적 변화가 필요함을 시사한다.

결국 AI 시대 교실의 성패는 기술 자체가 아닌 교육적 접근에 달려 있으며, AI를 잘 활용하는 교사가 그렇지 못한 교사보다 더 큰 교육적 효과를 창출할 수 있을 것이다.

생성형 AI가 불러올 교육 격차와 대응 방안

기술의 발전은 항상 새로운 형태의 사회적 격차를 만들어왔다. 생성형 AI도 예외가 아니다. 이미 우리 교실에 들어온 AI가 모든 학생에게 동등한 혜택을 주고 있을까? 아쉽게도 그렇지 않다. AI로 인한 교육 격차는 단순히 컴퓨터나 태블릿을 가지고 있느냐의 문제를 넘어선다.

가장 먼저 눈에 띄는 것은 AI 접근성 차이다. 좋은 성능의 AI 서비스들은 대부분 돈을 내야 사용할 수 있다. ChatGPT, Claude, Gemini와 같은 AI 서비스에서 제공하는 고급 모델을 활용하기 위해서는 월 구독료가 필요하고, 이것이 실제로 부담으로 작용할 수 있다. 또한 인터넷 환경이나 디지털 기기 보유 여부도 지역과 가정 형편에 따라 크게 차이가 난다. 이런 물리적 접근성 차이는 곧바로 학습 기회의 불평등으로 이어진다.

하지만 더 중요한 격차는 AI를 얼마나 잘 활용하는지에 있다. 같은 AI 도구를 사용하더라도, 이를 효과적으로 활용하는 능력은 학생마다 크게 다르다. AI에게 좋은 질문을 하는 방법, AI가 제공한 정보를 비판적으로 평가하는 능력, 여러 출처의 정보를 모아 자신만의 생각으로 재구성하는 능력은 저절로 생기지 않는다. 이런 능력은 체계적인 교육을 통해 길러져야 한다.

이런 격차를 줄이려면 국가와 사회가 함께 노력해야 한다. 먼저 AI 리터러시 교육을 정규 교육과정에 포함시켜야 한다. 모든 학생이 학교에서 체계적으로 기본적 AI 활용 능력을 배울 수 있어야 한다. 이는 단순한 기술 교육이 아니라, 비판적 사고력, 정보 활용 능력, 윤리적 판단력을 함께 기르는 교육이 되어야 한다.

특히 교육 취약계층 학생들을 위한 특별 프로그램도 필요하다. 방과 후 AI 교육, 지역 도서관이나 청소년 센터에서의 AI 수업 등을 통해 가정에서 지원받기 어려운 학생들에게도 충분한 학습 기회를 제공해야 한다.



생성형 AI는 교육에 혁신적 변화를 가져올 잠재력이 충분하다. 하지만 이런 변화가 모든 학생에게 공평하게 혜택을 주도록 하는 것이 교육의 공공성 측면에서 중요한 과제다. 기술 발전이 교육 불평등을 더 키우는 것이 아니라, 오히려 이를 줄이는 방향으로 나아가야 한다. 그래야만 AI가 진정으로 교육의 미래를 밝게 할 수 있을 것이다.

학교 통계 교육에서 데이터 윤리

권나영 | 인하대학교 수학교육과 교수



우리나라 수학과 교육과정은 학교 수학을 몇 개의 영역으로 나누고 있다.
통계와 관련된 영역은 영역명에서 꾸준히 변화해 왔는데, 2009 개정까지는 학교급에 상관없이 '확률과 통계'라는 영역명을 유지하다가 2015 개정에서부터 초등학교에서는 '자료와 가능성'이라는 이름으로, 중·고등학교는 '확률과 통계'라는 영역으로 바뀌었다.
그리고 최근 개정된 2022 개정 교육과정부터 초·중학교에서는 '자료와 가능성'이란 영역명을 사용해 초등학교와 중학교의 연계성을 강화하고, 고등학교는 선택과목명으로 '확률과 통계'를 사용한다. 영역명이나 과목명 및 학교 교육에서 가르치는 내용과 범위의 차이는 사회 및 교육 환경 변화에 따른 교육과정 설계 변경에 따른다.

그렇다면 학교 수학에서 통계와 관련된 영역은 왜 꼭 필요할까?

21세기는 디지털화가 가속되면서 데이터 수집 및 저장, 분석 수요가 많아지고 이에 따라 사회를 살아가는데 필요한 통계와 관련된 소양을 기르는 것이 중요해졌다. 학교 수학에서는 학년에 따라 통계 관련 개념 등의 내용을 배우고 자료 수집 및 분석 관련 활동을 한다. 이 활동을 통해 학생들은 통계적 소양을 기를 수 있다.
통계적 소양은 명확히 정의된 연구는 없지만, 대체로 자료를 조직하고 표를 작성하고 자료를 다양한 표현으로 나타내는 기능들로 이루어져 있고, 통계적 정보나 연구 결과를 이해하는데 이용된다고 한다.
통계적 소양은 단순히 통계적 지식이나 기술 습득에 그치는 것이 아니라, 실생활에서 접하는 통계 정보나 자료, 현상들을 해석하고 비판적으로 평가하고 이렇게 정리된 정보로 의사소통을 하는 것 등 사회로 나아간 이후의 삶과도 관련된다.

우리는 의사결정을 할 때 데이터에 의존한다. 그런데 데이터가 왜곡되어 있어 결정에 문제가 생길 수도 있고 데이터를 처리하는 능력이 부족할 수도 있다.
데이터를 분석하는 과정을 컴퓨터나 공학 도구에게 맡겨서 처리 능력을 해결할 수도 있지만, 그렇다고 해석상의 오류나 의도적인 결과 해석으로 인해 발생하는 통계 윤리의 문제는 여전히 남아있다. 이렇게 데이터를 다루는 일에서 발생하는 문제점들을 데이터에 대한 윤리 이슈라고 한다.
구체적으로는 데이터 생성에서부터 활용에 이르는 전 과정에서 개인의 사생활을 보호하고, 데이터의 신뢰성과 투명성을 유지하며, 사회적 책임을 다하기 위해 필요한 기준을 데이터 윤리라고 한다. 지금부터는 학교 수학에서 통계와 관련된 활동을 할 때 지켜야 할 데이터 윤리에 대해 논의하려고 한다.
학교 수학에서 데이터 윤리를 논의하기에 앞서 한 일화를 살펴보겠다.

2024년도 2학기에 학교현장실습학기제에 참여하는 한 예비교사의 수업을 참관한 적이 있다. 마침 중학교 1학년 자료의 정리와 해석 단원 마무리 시간으로 학생들이 만든 통계 포스터 발표가 있었다.
학생들은 조를 나누어 조사하고 싶은 주제별로 자료를 조사하고 그 결과를 통계 수업 시간에 배운 그래프나 그림을 이용해 분석 자료를 만들고 발표하는 시간이었다.
학생들은 개인적으로 궁금한 질문들(예를 들어, 점심 메뉴로 무엇을 선호하는지, 키가 얼마인지 등)을 반 학생들 또는 다른 반 친구들에게 물어보고 자료를 수집한 뒤, 막대 그래프, 원 그래프, 꺾은선 그래프 등 다양한 모양으로 포스터를 만들어 발표하였다.
조별 활동으로 진행되어, 조원들의 이지통계나 엑셀 등을 이용해 자료를 입력하고 원하는 모양으로 만들어내고 의미 있는 분석을 하는 등의 능력에서 차이도 있었고, ppt로 만드는 능력에서의 차이, 발표자의 발표 능력에 따른 전달력에서 차이가 있는 등 재미있는 수업이었다.



이 일화를 보면 최근 중학교에서 많이 하고 있는 통계 관련 발표대회나 포스터대회를 떠올릴 수 있다. 물론 수학 시간에 수행 평가로도 많이 진행되고 있다.

이런 활동을 통해, 예전 학교 수학 교육과정에는 통계 단원에서 이론적으로 자료를 표와 그래프로 표현하는 방법, 이를 분석하는 방법을 다루던 것과 달리, 최근에는 학생들이 직접 자료를 수집하고 그 자료를 표현하고 분석하는 과정을 하는 활동으로 바뀐 것을 알 수 있다.

실제로 1990년대 중반에 등장한 "빅데이터"라는 용어가 2000년대에 들어와서 학계와 산업계의 주목을 받기 시작하였고, 2010년 초반부터 교육 분야에서도 러닝 어낼러틱스(Learning Analytics)라는 개념과 함께 빅데이터가 주목받게 되면서 데이터와 관련된 통계 교육이 중요해지기 시작하였다.

우리나라는 2009 개정에서부터 최근 2022 개정 교육과정 문서에 이르기까지 통계 영역에 실생활 소재를 가지고 자료를 조사하고 표나 그래프로 표현하고 의미를 해석하는 내용을 강조하면서 실제 수업에서도 변화가 나타나기 시작하였다.

2009 개정 교육과정 이후부터 과정 중심 평가를 위한 수행평가의 한 예로 통계 포스터가 많이 활용되면서 그 변화는 더욱 두드러졌다. 이런 수업 활동들은 학교 교육과정에서 이론적이기만 하던 통계 교육이 실제 맥락에서의 통계 교육으로 바뀌고 있는 현실을 엿볼 수 있게 한다.

교사교육자로서 저자는 일화에 있던 수업을 참관하면서 예비교사가 실습 중 수업에서 학생들이 무슨 자료를 수집하게 하고 어떻게 분석하고 분석결과를 어떻게 표현하게 했는가 등에만 집중하였다. 그런데 여기서 한 걸음 더 나아가, 학교 수학에서 데이터 윤리를 함께 고려해야 한다고 주장하고 싶다. 다음 일화를 살펴보자.



2024년도 9월 즈음 아들이 아침에 학교 갈 준비를 하다 말고 갑자기 채팅방을 보고 있기에 뭘 하는지 물어봤더니 수업에서 하는 설문이라며 빨리 답해줘야 한다고 얼른 버튼을 눌렀다. 무슨 설문인지 물었더니 수학 시간에 발표를 해야 해서 조별로 질문을 만들고 반에서 운영하는 채팅방에 올려서 답을 받아서 정리해서 발표를 한다는 것이다.

보아하니 통계 시간에 발표를 위한 설문이었던 것 같다. 그런데 질문이 "너가 자주 사용하는 앱이 무엇이냐?"는 걸 물어보는 단 하나였다.

여기서 갑자기 의문이 들었다. 반 애들을 대상으로 질문을 해서 자료를 수집하는데, 응답을 받아도 되는지에 대한 동의 절차 같은 건 없을까 하는 거였다. 아들에게 물어봤더니 그런 건 선생님이 하라고 하지도 않았고 답만 받아서 숫자만 정리하면 된다고 하였다. 자료를 정리하고 발표를 하는데, 과연 숫자만 필요한 걸까?

이 글을 읽고 있는 독자들은 이 일화를 보면서 반 친구들을 대상으로 간단한 질문을 물어보고 결과를 정리해 발표하는데 무슨 데이터 윤리까지 생각해야 하는가라고 여길 수도 있다. 그런데 앞에서도 밝혔듯이 데이터 윤리는 데이터 생성에서부터 활용에 이르는 전 과정에서 필요한 기준이다.

학생들이 학교에서 하는 통계 수업에서의 설문조사는 사실 통계 업체에서 하는 설문조사와는 비교가 되지 않을 정도로 단순하고 짧은 질문 수준이라고 생각할 수 있다. 하지만 학생들은 이 과정에서 설문조사라는 개념을 알게 되고 이런 식으로 타인의 개인 정보를 수집하는 과정을 아주 간단한 방식으로 경험하게 된다. 이 첫 과정에 타인에게 개인적인 질문으로 정보를 수집하는 행위를 하기 전에 이 질문에 대한 응답을 수업 자료로 써도 되는지 동의를 얻는 과정이 필요하다는 것을 알려주어야 한다. 이것은 데이터 윤리 교육의 가장 첫 걸음이다.

우리는 학교 통계 교육에서 통계 전문가를 길러내려는 것이 아니라, 앞으로 사회에서 삶을 살아가기 위해 필요한 통계적 소양을 기르기 위해 쉽고 초보적인 방식으로 자료 수집과 분석의 과정을 경험하게 한다. 이 학생들이 나중에 커서 길거리에서 설문조사원을 만나 설문에 응답하게 될 수도 있고, 자신의 분야에서 필요한 설문조사를 자신이 만들게 될 수도 있다. 이럴 때 데이터 윤리를 알고 있다면 자신이 원하지 않는 설문조사에 응하지 않아도 된다는 사실을 알고 그에 대응할 수 있을 것이고 설문조사를 만들어야 한다면 이 때 필요한 원칙들을 지키기 위해 노력할 것이다.

데이터 윤리를 지키지 않은 자료 수집과 분석의 문제점을 알게 하는 가장 손쉬운 방법은 자료 수집과 분석을 처음 경험하는 학교 통계 교육 단계에서 언급하는 것이라고 생각한다. 이 단계는 처음 통계를 접하는 순간이기 때문에, 통계를 다루기 위해 필요한 내용과 기능에 대한 것뿐만 아니라 기초적인 윤리에 대해서도 함께 배우는 것이 중요하다.



디지털 기반 사회에서 빅데이터는 계속 생성되고 우리들은 매일 데이터가 관련된 환경에 노출될 것이다. 우리 자신이 관련된 데이터는 어떤 식으로든 생성되고 활용하고 활용당할 수 있음을 인식하는 것이 중요하다. 학교 통계 교육은 학생들에게 이러한 인식을 생성하게 하는 첫 경험을 주게 될 것이다.

지금 교사에게 요구하는 것은 교사가 통계 전문가가 되기를 바라는 것이 아니라, 통계 시간에 관련된 활동을 하면서 학생들에게 가장 기본적인 데이터 윤리에 대한 것도 함께 노출하기를 바라는 것이다. 예를 들면, 학생들에게 데이터 윤리의 정의에 대해 한 번 알려줄 수도 있고, 설문조사 등의 자료 수집 활동에서 유의점 등으로 설명할 수도 있다. 중요한 것은 데이터 윤리라는 것이 통계를 하는 과정에 중요한 원칙이라는 인식을 심어주는 것이다.

빅데이터를 생성하고 분석하는 것은 기계와 기술로 대체될 수 있을 지라도 그 기술로 발생할 수 있는 문제에 대응하기 위해 맥락과 상황에 맞는 윤리 원칙을 세우는 것은 인간의 역량이다. 인공지능과 디지털 사회가 발전될수록 데이터 윤리는 점점 더 정교화될 것이고 그 기준을 정립하고 활용하기 위해서는 학생 시절 통계 관련 활동에서 데이터 윤리에 대해 직접 경험하고 실천해 보는 것이 가장 좋은 교육이 될 것이다.



생성형 AI 시대의 필수 역량, 디지털 리터러시

김한성 | 고려사이버대학교 소프트웨어교육 교수

① 들어가며

"수업시간에 검색도 제대로 하지 못하는 경우가 많다는 점에서 에듀테크 교육을 하기 위해선 디지털 리터러시 교육이 우선돼야 한다고 생각해요[1]."와 같은 한 현직 교사의 이 말처럼, 겉보기에는 스마트폰과 태블릿을 능숙하게 다루는 오늘날 학생들도 정작 필요한 정보를 찾아 비판적으로 평가하고 자기에게 적절한 형태로 활용하는 기초적인 역량에서는 부족함을 보이곤 한다.

더욱이 ChatGPT의 등장 이후 글쓰기부터 코딩까지 생성형 AI가 학습 도구로 급부상하면서, 정보를 비판적이고 창의적으로 활용할 능력이 없다면 오히려 새로운 '디지털 문맹'이 될 수 있다는 우려도 나온다[2]. 실제로 교육 현장에선 AI 기반 디지털 교과서 도입이 현실이 되었고[3], 미래 사회의 핵심 역량으로 디지털 리터러시가 부각되고 있다.

본 고에서는 디지털 리터러시의 개념과 중요성을 살펴보고, 국내외 연구 동향과 한국 학생들의 수준을 통해 디지털 전환 시대의 기본 소양으로서 디지털 리터러시의 역할을 짚어보고자 한다.

② 디지털 리터러시의 개념과 국내외 연구 현황

디지털 리터러시(Digital Literacy)란 디지털 시대를 살아가는 데 개인에게 요구되는 핵심 역량을 의미한다. 1990년대 Gilster 등이 처음 사용한 이 개념은 당시에는 주로 정보 기술을 이해하고 활용하는 능력을 가리켰지만, 오늘날에는 그 의미가 크게 확장되었다고 할 수 있다[4]. 디지털 리터러시가 단순히 컴퓨터나 스마트기기 사용 능력에 그치지 않고, 디지털 기술을 활용한 의사소통과 협업 능력, 디지털 환경에 적응하는 문제해결 능력, 그리고 올바른 사용을 위한 윤리의식까지 포괄하는 필수 소양으로 자리매김한 것이다. 다시 말해, 디지털 리터러시는 정보를 찾아 평가하고 조합하여 새로운 지식을 만들어내는 능력부터 안전하고 책임감 있게 기술을 활용하는 태도까지 아우르는 광범위한 역량이라 할 수 있다.

국제적으로도 디지털 리터러시의 중요성에 대한 인식이 높아지면서 여러 연구와 지표가 개발되었다. 예를 들어 OECD 교육 2030에서는 미래 학습자에게 가장 중요한 역량으로 디지털 리터러시와 데이터 분석력을 제안하였다[5]. 이는 전통적으로 교육계에서 중시되어온 창의성이나 문제해결력을 앞서는 순위로, 디지털 사회에서 정보 활용 능력과 데이터 문해력이 얼마나 중요한지 보여준다.

유럽연합(EU) 또한 디지털 역량 프레임워크(Dig Comp)를 통해 정보·데이터 활용, 커뮤니케이션, 콘텐츠 제작, 안전, 문제해결의 다섯 영역으로 시민의 디지털 역량을 정의하고 있으며, UNESCO 역시 미디어·정보 리터러시교육을 강조하며 글로벌 차원의 노력을 이어가고 있다. UNESCO는 최근 전통적으로 강조해오던 미디어, 정보 리터러시와 더불어 생성형 AI에 대한 이해와 올바른 활용에 대해서도 관심을 기울여야함을 제안하고 있다[6].

한편 국내에서는 한국교육학술정보원(KERIS)을 중심으로 초등, 중등 학생 대상의 국가 수준 디지털 리터러시 수준을 측정하고 시사점을 제시하고 있다. 여기서 국가 수준이라 함은 전국에 있는 초등학생과 중학생을 대상으로 한 전수조사가 아니라, 이를 대표할 수 있는 체계적인 표집을 통해 전체 학생들의 수준을 가늠할 수 있는 형태로 수행되는 연구를 의미한다.

2007년부터 시작된 국가수준 디지털 리터러시 수준 측정연구는 초기 선다형 검사에서 시작해, 2023년부터 수행형 검사로 전환되며 학생들이 "얼마나 잘

1) https://it.chosun.com/news/articleView.html?idxno=2023092109953
2) https://shorturl.at/5cWvr
3) https://www.keris.or.kr/main/aidtMain.do
4) 김한성 외(2024). 2024년 학생 디지털 리터러시 수준 측정 및 발전방안 연구. KERIS.
5) https://www.oecd.org/en/about/projects/future-of-education-and-skills-2030.html
6) Stefania Giannini(2023) Generative AI and the future of education. UNESCO.

알고 있는가?"에서 "얼마나 잘 해결할 수 있는가?"의 관점으로 검사를 재설계하였다. 아울러, 2022 개정 교육과정에서부터 강조된 인공지능 관련 내용도 포함해 디지털에 대한 이해와, 디지털 기반의 문제해결 능력, 코딩에 대한 기초 이해와 응용, 그리고 인공지능에 대한 이해까지 확장되었다고 할 수 있다[7].

③ 한국 학생의 디지털 리터러시 수준[8]

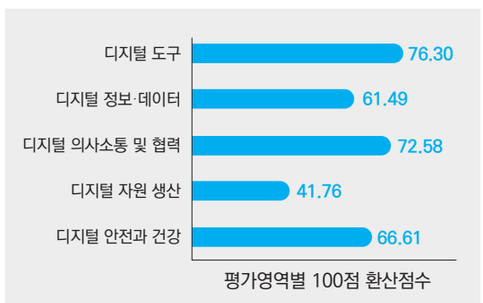
그러면, 우리나라 초·중학생들의 디지털 리터러시 수준은 어느 정도일까? 2024년 KERIS에서 실시한 국가 수준 디지털 리터러시 검사결과에 따르면, 한국 학생들의 디지털 리터러시에는 뚜렷한 강점과 약점이 혼재하는 것으로 나타났다.

우선 대부분의 학생들은 기본적인 디지털 도구 활용 능력과 안전한 사용에 관한 지식에서 비교적 높은 성취를 보였다. 초등학생의 경우 디지털 기기 다루기(디지털 도구) 영역의 점수가 가장 높았고, 중학생의 경우 디지털 안전과 건강영역의 점수가 가장 높게 나타나 전반적으로 도구 활용 능력이나 사이버 안전 의식은 양호한 수준임을 알 수 있었다.

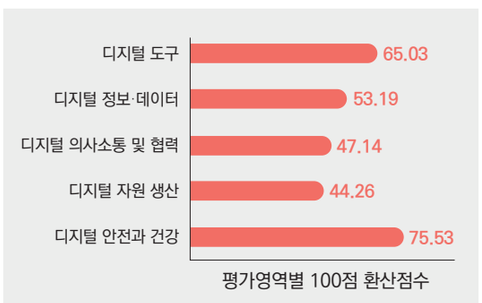
반면, 디지털 자원 생산 영역, 즉 새로운 디지털 콘텐츠를 창의적으로 제작하는 능력은 초등학생과 중학생 모두에서 다른 영역에 비해 낮은 성취 수준을 보였다. 물론 해당 영역의 문항은 상대적으로 높은 인지적 요구를 포함하고 있어 단순 비교에는 주의가 필요하다.

그럼에도 불구하고, 디지털 사회에서 정보의 단순 소비를 넘어 새로운 가치를 창출하는 능력은 점차 핵심 역량으로 부상하고 있으며, 이를 고려할 때 해당 영역의 교육적 중요성은 더욱 강조되어야 한다.

학생들의 성취 수준 분포를 살펴보면 새로운 시사점을 얻을 수 있다. 검사 결과에 따르면 학생들의 디지털 리터러시 수준은, 가장 높은 4수준(최고)부터 가장 낮은 1수준(최저)까지 분류했을 때, 초등학생들은 3수준과 4수준에 대부분 분포하고 1·2수준에 해당하는



[그림 1] 영역별 100점 환산 점수(초등학생)



[그림 2] 영역별 100점 환산 점수(중학생)



비율은 낮게 나타났다. 이에 대한 해석은 다양할 수 있지만, 디지털 리터러시 검사가 말 그대로 기본 소양을 측정하는 관점에서 개발된 도구인 만큼 1, 2수준에 해당하는 학생의 경우 디지털 리터러시 수준이 현저히 낮은 학생으로 볼 수 있을 것이다.

중학생들은 초등학생과 조금 다른 양상을 보였다. 높은 수준인 3·4수준의 비율이 높긴 하지만, 초등에 비해 최하위인 1수준 학생들의 비율이 2수준 학생들의 비율보다 높게 나타난 것이다. 이는 일부 중학생들의 경우 디지털 기초 역량에서 매우 어려움을 겪는 집단이 존재함을 보여준다. 이를 통해 학생들 간의 디지털 리터러시 격차가 발생할 수 있을 가능성과 함께, 초등학교에서 중학교로 넘어가는 시기에 디지털 리터러시 격차가 벌어지지 않도록 지속적인 지원이 필요성을 보여주기도 한다.

2024년 연구에서는 2023년과 2024년 결과를 비교해 학생들의 수준이 어떻게 변화했는지도 관찰하였다. 그 결과, 2024년 학생들의 디지털 리터러시 수준이 전반적으로 향상된 것으로 나타났다. 초등학생의 경우 2024년 평균 점수가 2023년에 비해 유의미하게 높아졌고, 성취수준 분포에서도 상위권(4수준, 3수준) 비율이 증가하며 하위권(2수준, 1수준) 비율이 줄어든 것으로 나타났다. 중학생 역시 2024년 평균 점수가 전년도보다 높아졌고, 특히 최고 수준인 4수준에 속하는 학생 비율이 약 12%p 증가하여 전체적으로 중상위 집단이 크게 늘어난 모습을 보였다. 하지만, 중학생의 최하위 집단(1수준) 비율도 소폭 증가한 점은 앞으로 어떠한 양상으로 변화할지 추이를 주의 깊게 살펴봐야 할 부분이기도 하다.

한편, 성별에 따른 분석에서는 여학생이 남학생보다 디지털 리터러시 점수가 전반적으로 높게 나타나는 경향이 이번 2024년 조사에서도 확인되었다. 이 같은 성별 차이는 2007년 첫 국가수준 디지털 리터러시 수준 측정 이후 꾸준히 관찰된 현상이며, 2018년 컴퓨팅 사고력 요소가 새롭게 포함되었던 해에는 일시적으로 남학생이 일부 영역에서 우세를 보이기도 했다. 그러나 2024년에는 모든 영역에서 여학생이 남학생보다 높은 성취 수준을 보였다.

이러한 결과는 디지털 리터러시가 단순한 기술 숙련도 이상의 역량, 즉 계획성과 집중력, 문제 해결을 위한 전략적 접근, 정보 판단력 등 복합적 사고력 및 학습 습관과 밀접하게 관련된 소양임을 시사한다.

또한 성별에 따라 디지털 학습 환경에서 보이는 참여 방식과 동기, 자기조절 능력의 차이를 보다 세밀하게 이해

7) 김한성 외(2024). 2024년 학생 디지털 리터러시 수준 측정 및 발전방안 연구. KERIS.

8) 3장의 내용은 '김한성 외(2024). 2024년 학생 디지털 리터러시 수준 측정 및 발전방안 연구. KERIS.'의 연구의 주요 결과를 요약하였음.

하고, 이를 기반으로 학생 맞춤형 교육 전략을 수립하는 것이 필요함을 보여준다. 특히 남학생의 몰입도나 지속적인 문제해결 전략을 촉진할 수 있는 학습 설계와 환경 지원이 병행되어야, 디지털 리터러시의 성별 편차를 줄이고 모두에게 효과적인 역량 함양이 가능할 것이다.

④ 나가며

디지털 리터러시는 이제 읽기, 쓰기, 셈하기와 같은 기본 문해력에 비견될 정도로 중요한 시대적 필수 역량이 되었다. 특히 AI와 데이터 기술이 교육과 삶 전반을 변화시키는 디지털 전환 시대에는, 디지털 리터러시가 곧 미래 사회를 살아갈 학습자의 생존 능력이라고 해도 과언이 아니라 할 수 있다. 그렇다면 앞으로 어떤 역량을 중심으로 디지털 리터러시 교육을 강화해야 할까?

디지털 리터러시는 앞으로 새로운 디지털 사회에 적응하고, 그 속에서 능동적으로 살아가기 위한 기초 소양으로 지속적으로 확장될 것이다. 특히 생성형 AI의 부상은 단순한 기술 활용 능력을 넘어, AI가 생성한 정보를 비판적으로 평가하고, 이를 바탕으로 독창적인 결과물을 만들어내는 능력을 요구하고 있다. 즉, 이제는 AI에 대한 이해, 윤리적 판단, 전략적 활용 역량이 디지털 리터러시의 핵심 요소로 포함되어야 한다.



그러나 아직까지 우리는 언제부터, 어떤 방식으로 학생들에게 생성형 AI를 교육해야 하는지에 대한 충분한 연구와 합의가 이루어지지 않은 상태이다. AI 도구의 무분별한 활용은 학생들의 사고력 저하와 학습의 의미를 약화시킬 우려가 있으며, 교육적 기회로 전환하기 위해서는 단계별 설계와 지도 원칙이 뒷받침되어야한다.

따라서 생성형 AI를 교육 현장에 도입하기 위해서는 연령별 발달 특성과 학습 환경을 고려한 체계적 연구와 실증 기반의 정책 수립이 선행되어야 하며, 그 이후에야 학생들의 성장에 도움이 되는 방향으로 교육에 통합될 수 있을 것이다.

이러한 점에서, 생성형 AI를 포함한 미래형 디지털 리터러시 교육은 충분한 논의와 준비 위에 설계되어야 하며, 학교 교육과정과 교수·학습 방식의 변화가 이를 뒷받침해야 할 것이다. 다행히 우리 교육 현장에서도 이러한 변화의 흐름이 서서히 나타나고 있다. 예를 들어, 일부 학교에서는 프로젝트 기반 수업을 통해 학생들이 스스로 자료를 탐색하고, 디지털 도구를 활용해 결과물을 창작하는 활동이 확산되고 있으며, 코딩 수업



에서도 문법 중심의 기술 습득을 넘어 논리적 사고와 문제 해결 능력을 강조하는 방향으로 변화하고 있다. 최근에는 생성형 AI 도구를 활용한 탐색적 학습 활동도 시도되고 있으며, AI 리터러시의 요소를 자연스럽게 접목하려는 시도도 증가하고 있다[9].

더불어, 디지털 시민성 교육 자료를 개발하여 초·중등 학생들이 윤리적이고 책임감 있는 태도로 디지털 세계를 접할 수 있도록 돕는 노력도 병행되고 있다[10]. 정책 차원에서는 교사 대상 연수 프로그램을 통해 생성형 AI 시대에 요구되는 새로운 문해력을 함양할 수 있도록 지원하고, 교육 자료와 평가 도구의 지속적 개발을 통해 학교 현장의 실천을 적극적으로 뒷받침해야 할 것이다.

이처럼 학교 현장과 정책의 변화가 함께 이뤄진다면, 우리는 비로소 생성형 AI 시대에 걸맞은 디지털 리터러시 교육의 기반을 마련할 수 있을 것이다. 디지털 리터러시 없이는 더 이상 미래를 논하기 어렵다는 말이 나올 정도로[11], 지금의 학생들에게 디지털 리터러시는 더 이상 선택이 아닌 삶의 필수 기술이 되었다. 이 중요한 역량을 인식하고 체계적으로 키워나간다면, 우리 아이들은 정보의 홍수와 AI 기반 도구들 속에서도 주도적으로 문제를 해결하고, 창의적으로 가치를 창출하는 인재로 성장할 수 있을 것이다. 디지털 전환의 거센 파도 속에서 '디지털 리터러시'라는 나침반을 손에 쥔 미래 세대가, 단순한 기술 사용자에 그치지 않고 더 나은 시민으로서 세상을 이끌어가는 주체로 성장하길 기대한다.

9) https://keris.or.kr/main/ad/pblcte/selectPblcteETCInfo.do?mi=1142&pblcteSeq=13792
10) https://xn--2z1b40gs9nlqcf0n.kr/front/archive/archiveDetail.do?searchArchiveId=ARC000000000553
11) https://www.esgkedu.com/news/articleView.html?idxno=312

통계로 바라보는 세상이야기

신동헌 | 도서출판 지일북스 대표

75명 사상자, 완전 복구에 100년

지난 3월 21일 경북 의성에서 시작한 산불로 역대 최대 규모의 피해가 발생한 것으로 추정됩니다. 인명피해는 사망자 30명을 포함하여 총 75명의 사상자가 발생하였으며 산불영향구역은 총 4만 8천여 ha로 여의도 면적의 50배에 달하며, 주택 3천여 동이 전소되고, 국가유산 피해 30건과 농업시설 2천여건 등 시설 피해도 상당합니다.

최근 10년간(2014~2023년) 우리나라에서는 5,667건의 산불이 발생하여 산림 40,073ha와 주택·건축물 등에 피해를 입었다고 하는데요. 이는 여의도 전체 면적의 약 138배에 해당하는 규모라고 합니다. 대형 산불이 일어나 피해를 입은 지역이 다시 나무를 키울 수 있게 되는데는 2~4년의 시간이 걸리며, 생태계가 원상 복원되기까지는 그보다 훨씬 긴 40~100년이 걸린다고 합니다.

대한민국을 이끌어갈 청년들에게

통계청이 발표한 '2024년 5월 경제활동인구조사 청년층 부가조사' 결과에 따르면 청년층(15~29세)이 졸업 후 첫 일자리를 구하는데 11.5개월 소요되는 것으로 나타났는데요. 이는 지난해보다 1.1개월 늘어난 것으로 2006년 관련 통계 작성 이래 역대 최장기간이라고 합니다. 고물가, 경기침체 등으로 일자리가 줄어들면서 청년들이 취업에 걸리는 시간이 늘어난 것으로 보입니다.

그렇다면 청년층의 첫 일자리 평균 근속 기간은 얼마나 될까요? 2024년 5월 기준 첫 직장 평균 근속 기간은 1년 7.2개월로 전년 동월 대비 0.6개월 증가했는데요. 청년층들이 첫 일자리를 그만둔 이유는 보수나 근로 시간과 같은 근로 여건 불만족이 가장 컸습니다. 2024년 7월 기준 일도 구직활동도 하지 않고 그냥 '쉰다'는 청년들이 7월 기준 역대 최대치를 기록했다는 뉴스였습니다.

생성형 AI, 열풍 넘어 일상으로

최근 챗GPT 열풍이 뜨겁습니다. 업무를 위한 기획서를 작성하거나, 사진 이미지를 애니메이션 풍으로 바꾸고, 심지어 AI와 연애에 빠지는 등 생성형 AI는 우리 일상에 깊숙하게 들어온 듯 합니다. 경향신문에 따르면 GPT를 활용해 기존 사진을 일본의 애니메이션 제작사 스튜디오 지브리의 스타일로 모방하는 밈이 확산하면서, 챗GPT의 일간 이용자 수가 처음으로 120만명대를 기록한 것으로 나타났다고 보도했습니다. 지난 4월 1일 데이터 플랫폼 기업 아이지에 이웍스의 모바일인덱스에 따르면 지난달 27일 기준, 챗GPT 국내 일간 활성 이용자 수(DAU)는 역대 최다인 125만 2925명으로 집계됐고, 지난 달 10일 챗GPT DAU는 103만3733명으로, 첫 100만명대를 기록했는데 약 2주 만에 최다 기록을 다시 경신한 셈입니다.

제철코어, 흐려진 사계절에 대한 향수

요즘 10대, 20대 사이에서 '힙한 라이프스타일'로 떠오르고 있는 트렌드가 있습니다. 바로 '제철코어'입니다. '제철코어'는 부모님이 과일 농사를 짓거나, 맛있는 제철 과일을 풍족하게 즐기는 사람이 새로운 부러움의 대상이 되면서 생긴 개념인데요. 그런데, 단순히 제철 과일을 많이 먹을 수 있어서가 아니라, 특정 계절과 감성을 즐기고 싶은 Z세대의 감성이 반영된 트렌드라는 점이 흥미롭습니다. Z세대가 '제철코어'에 열광하는 이유는 사라져가는 사계절에 대한 향수 때문입니다.

통계청의 「2024년 사회조사 보도자료」에 따르면 환경 문제 중 기후변화에 불안을 느끼는 비율이 53.2%였으며, 기상청의 「기후위기 감시 및 예측 서비스 발굴을 위한 설문조사 결과」에서도 현재 대한민국이 기후위기에 직면해 있다고 답한 비율이 89.9%로 나타났습니다.

4월 22일 지구의 날, 자연과 인간 공존

1970년 4월 22일 미국 위스콘신주의 게이로드 넬슨 상원의원이 앞서 1969년 미국 캘리포니아주에서 발생한 해상원유 유출사고를 계기로 환경문제에 관한 범국민적 관심을 불러일으키기 위해 '지구의 날'을 주창한 것이 계기가 되었는데, 유엔이 정한 세계환경의 날(6월 5일)과는 달리 순수 민간운동에서 출발하였습니다.

기상청에서는 지난 2024년 여름철(6~8월) 기후 특성에 대해 높은 기온과 열대야 기승, 장마철 집중호우가 특징이었다고 발표했고, 환경부가 국민 1천 501명과 전문가 504명을 대상으로 실시 한 「2023년 환경보전에 관한 국민 의식 조사」 결과, 지난 6~8월 전국 평균기온은 25.6도로 평년 23.7도보다 1.9도나 높았으며, 1973년 이후 가장 높았던 기온이라고 하는데요. 전국 평균 폭염일수 또한 24.0일로 역대 3위를 기록, 열대야 일수는 20.2일로 역대 1위를 기록하며 우리 국민들이 많은 불편과 피해를 겪었습니다.

국민 삶의 질 2024, 우리 국민의 삶은?

2000년대 이후부터 국제기구를 중심으로 삶의 질 측정에 대한 중요성이 부각되어 왔는데요, 우리나라 사람들은 연령이 높아질수록 건강 문제나 경제적 부담, 사회적 고립 등이 삶의 질에 부정적인 영향을 미칠 가능성이 있습니다. 2023년과 2024년 삶의 만족도는 10점 만점에 6.4점으로 2022년 대비 0.1점 감소했고, 약간의 감소폭을 보이긴 했지만, 2021년과 2022년의 증가폭에 빗대어 봤을 때 큰 하락세를 보인 것은 아닙니다. 이는 전반적인 사회·경제적인 환경이 안정적인 수준을 유지하고 있음을 보여주며, 남자와 여자 삶의 만족도 모두 6.4점으로 성별에 따른 차이가 없었습니다. 이는 남녀 모두 비슷한 사회적·경제적 환경에서 살아가며, 삶의 질에 대한 평가가 크게 다르지 않다는 점을 보여주는데요. 30대 이하의 만족도는 6.5~6.6점, 40~50대는 6.4점, 60세 이상은 6.3점으로 연령이 높아질수록 만족도가 낮게 나타났습니다.

당신의 건강은 안녕하신가요?

통계청이 발표한 「2023년 사망원인 통계 결과」에 따르면 2023년 사망자 수는 353천 명으로 전년 대비 20천 명(-5.5%) 감소했는데요. 연령별로 보면 80세 이상의 사망자가 전체 사망자의 54.0%로, 10년 전에 비해 16.7%p, 전년과 비교하면 0.2%p 증가한 것으로 나타났습니다.

2023년 상위 10순위 사망원인은 악성신생물(암), 심장 질환, 폐렴, 뇌혈관 질환, 고의적 자해, 알츠하이머병, 당뇨병, 고혈압성 질환, 패혈증, 코로나19였습니다.

특히 우리 국민의 3대 사망원인은 암, 심장 질환, 폐렴으로 전체 사망의 41.9%를 차지하였는데요. 사망의 외인 사망률은 자살(27.3명). 운수사고(6.4명), 추락사고(4.9명) 순으로 높았으며, 전년 대비 화재사고(-16.8%), 중독사고(-11.4%), 추락사고(-6.9%) 등의 사망률은 감소, 익사사고(9.3%), 자살(8.5%) 사망률은 증가하였습니다.

늘어나는 어린이 미디어 이용시간

한국언론진흥재단이 발표한 「2023년 어린이 미디어 이용 조사」 결과에 따르면 2023년 우리나라 어린이의 하루 평균 미디어 이용 시간은 세계보건기구(WHO) 권고 기준의 약 3배에 달하는 3시간 6분으로 나타났습니다.

2020년부터 전국에 거주하는 만 3세 이상 만 9세 이하 어린이의 보호자를 대상으로 어린이의 '어린이 미디어 이용 조사'를 매년 실시하고 있는데요. 우리나라 어린이(만3세~만9세)들이 가장 많이 사용하는 미디어 기기는 스마트폰(77.6%)이었으며 인터넷 기능이 있는 스마트 TV(65.6%), 태블릿PC(57.1%), 컴퓨터(24.7%), 인터넷 기능이 없는 일반 TV(17.1%) 등이 그 뒤를 이었습니다.

어린이들의 미디어 기기별 하루 평균 이용 시간은 텔레비전이 1시간 13분으로 가장 길었고, 스마트폰이 1시간 3분, 태블릿PC 38분, 컴퓨터 12분 순으로 나타났습니다.

청소년 건강, 오늘 습관이 내일 결정



질병관리청에서 발표한 「청소년건강행태조사」 결과를 보면 청소년의 주 5일 이상 아침식사 결식률을 살펴보면 2021년 38.0%, 2022년 39.0%, 2023년에는 41.1%로 나타났는데요. 신체활동을 실천하는 청소년의 비중은 어느 정도일까요? 2023년 주 3일 이상 고강도 신체활동을 실천하는 청소년의 비중은 41.2%로 전년 대비 4.1%p 증가한 것으로 나타났습니다.

학교급별 더 자세하게 살펴보면, 2023년 중학생 중 주 3일 이상 고강도 신체활동을 실천하는 청소년의 비중은 51.2%로 전년대비 5.1%p 증가했으며, 고등학생의 경우 30.8%로 전년 대비 3.3%p 증가한 것으로 나타났습니다. 세계보건기구(WHO)에 따르면 청소년기에는 주 2~4번 운동하며 3회 이상 근육운동을 병행할 것을 권고하고 있을 만큼 청소년에게 규칙적인 신체활동은 중요한 건강관리 방법 중 하나입니다.

장래인구추계로 보는 2072년 지구촌

2024년 전 세계 인구는 81억 명을 돌파했고 계속 증가할 것으로 전망하고 있습니다. 통계청이 발표한 「2022년 기준 장래인구추계를 반영한 세계와 한국의 인구 현황 및 전망」 결과에 따르면 2024년 세계 인구는 81억 6천만 명으로 1970년과 비교해 2.2배 수준으로 증가, 2072년에는 102억 2천만 명까지 증가할 것으로 전망되었는데요. 반면, 2024년 한국 인구는 5천 2백만 명에서 2072년에는 3천 6백만 명으로 감소할 것으로 전망되었습니다.

2072년 대륙별 인구 전망은 대부분 2024년 대비 증가할 것으로 보이는데요. 한국이 속한 아시아 인구 역시 2024년 48억 1천만 명에서 2072년 51억 5천만 명으로 1.1배 수준 증가할 것으로 전망되었는데요. 반면, 유럽 인구는 2024년 7억 5천만 명에서 2072년 6억 4천만 명으로 0.9배 수준으로 감소할 것으로 전망되었습니다.

새 보금자리에서 첫걸음, 국내인구이동

통계청이 발표한 「2024년 국내인구이동통계」 결과에 따르면 전국 이동자 수는 628만 3천 명으로, 전년과 비교했을 때 15만 5천 명이 증가했습니다. 인구 100명당 이동자 수를 나타내는 인구이동률 역시 전년대비 0.3%p 증가한 12.3%로 나타났습니다. 전국 이동자 628만 3천 명 중 282만 6천 명이 20~30대였는데요, 이는 전체 이동률의 44.9%를 차지하는 높은 수치입니다.

젊은 층의 인구이동이 활발하게 일어나는 데 비해 60대 이상 연령대에서는 인구이동이 낮았고, 연령별 이동률 성비(여자 1백 명당 남자)는 40대가 109.5명으로 가장 높게 나타났으며, 80세 이상 연령대에서 74.0명으로 가장 낮았습니다. 전입자가 전출자보다 많아 순유입률이 높게 나타난 곳은 인천(0.9%), 세종(0.7%), 충남(0.7%)이었고, 광주(-0.6%), 제주(-0.5%), 서울(-0.5%)은 순유출률이 높았습니다.

인생 후반전을 준비하는 고령층

통계청이 발표한 「경제활동인구조사」 결과에 따르면 2023년 60세 이상 취업자 수는 622.3만 명으로 전년 대비 36.6만 명 증가했는데요. 2024년 9월 경제활동인구조사에서는 60세 이상 고령층 취업자가 전체 연령대에서 1위를 차지, 1982년 관련 통계 작성 이후 60대 취업자 수가 50대 취업자 수를 처음으로 제친 것으로 나타났습니다.

고용노동부가 발표한 「고용형태별 근로실태조사」에 따르면 근로자의 평균 연령은 2023년 43.8세로 매년 증가하였습니다. 이는 근로자의 고령화가 점차 심화되고 있음을 보여주는 지표로 앞으로도 근로 현장의 고령화는 심화될 것으로 예상됩니다.

2024년 5월 고령층(55~79세) 취업자의 산업별 비중을 보면 보건·사회복지(12.8%), 농림어업(12.4%)에서 높았으며, 예술·스포츠·여가(1.1%), 금융·보험업(1.9%)에서 낮게 나타났습니다.

건강한 삶을 위한 저속노화 실천



최근 들어 건강 산업이 성장하고, 건강한 라이프스타일에 관한 관심이 높아지면서 '저속노화'가 하나의 트렌드로 자리 잡고 있는데요. 단순히 수명을 연장하는 것이 아니라, 삶의 질을 높이는 핵심 요소가 바로 '저속노화'입니다. 통계청 「장래인구추계」에 따르면 2025년 대한민국의 기대수명은 84.5세로 예상됩니다. 기대수명이 늘어난 만큼 노후기간이 같이 길어지면서 건강하고 활기찬 삶을 유지하는 것이 중요합니다. 특히 65세 이상 인구가 꾸준히 증가하는 가운데, 2024년에 994만 명이던 65세 이상 인구는 2028년에는 1,213만 명에 이를 것으로 전망되는데요. 잡곡과 두류가 '저속노화'에 효과적이라는 점이 알려지면서 2022년 0.9kg이었던 1인당 연간 잡곡 소비량은 2024년 1.4kg, 두류 소비량은 2022년 1.7kg에서 2024년 1.9kg으로 증가한 것으로 나타났습니다.

연금, 안정적인 노후를 위한 디딤돌

통계청의 「연금통계」에 따라 연금 수급자 수와 수급률을 살펴보면 통계 작성 최초 시점인 2016년부터 연금을 받는 수급자가 지속적으로 증가하는 것으로 나타났습니다.

2022년의 연금 수급자는 818만 2천 명(수급률 90.4%)으로, 2016년 이후 지속적으로 수급자, 수급률 모두 증가했습니다. 65세 이상 연금 수급자의 월평균 연금 수급금액은 지속적으로 증가하는 것으로 나타났습니다.

2016년 월평균 42만 원이었던 수급금액이 2022년에는 65만 원으로 상승했는데요. 기초연금 및 국민연금 제도 변경 등으로 그 수준에서 약간의 차이는 있지만, 전년대비 평균적으로 5~10% 지속적으로 상승하고 있는 것으로 나타났습니다. 2022년 기준으로 연금 수급자의 수급금액 분포를 살펴보면, 월평균 연금 수급금액이 약 65만 원, 수급금액의 중위수는 약 42만 원으로 나타났습니다.

살고 일하고 쉬는, 새로운 농촌

올해 3월, 정부에서는 농촌에 사람·일자리를 유인하고 농촌정책을 전환하기 위해 '농촌소멸 대응 추진전략'을 수립·발표하였는데요. 농업인의 날은 한자 十一(11)을 합치면 十月十月이 되고 흙 토(土)는 농업의 터전이 되는 것에 착안해, 농업의 중요성을 되새기고자 정부가 1996년에 지정한 법정 기념일입니다.

가래떡 데이는 농업인의 날을 널리 알리고 쌀 소비 촉진을 위해 이로부터 10년 뒤인 2006년에 지정되었습니다.

통계청의 「2023년 농림어업조사 결과」에 따르면 2023년 12월 1일 기준 현재 우리나라 농가는 99만 9천 가구로 고령에 따른 농업 포기, 전업 등으로 전년대비 2만 4천 가구 감소하였는데요, 처음으로 100만 가구 밑으로 떨어졌고, 농가 인구 역시 전년대비 7만 7천 명 감소한 208만 9천 명으로 나타나는 등 농촌소멸도 빠르게 진행될 전망입니다.

공경과 예의 중시하는 우리말

국립국어원에서 만 20세 이상~만 69세 이하 성인 남녀 5,000명을 대상으로 한 「2020년 국민의 언어 의식 조사」 결과를 보면 국민들의 73.1%가 높임말 사용에 대해 긍정적으로 생각하고 있는 것으로 나타났습니다.

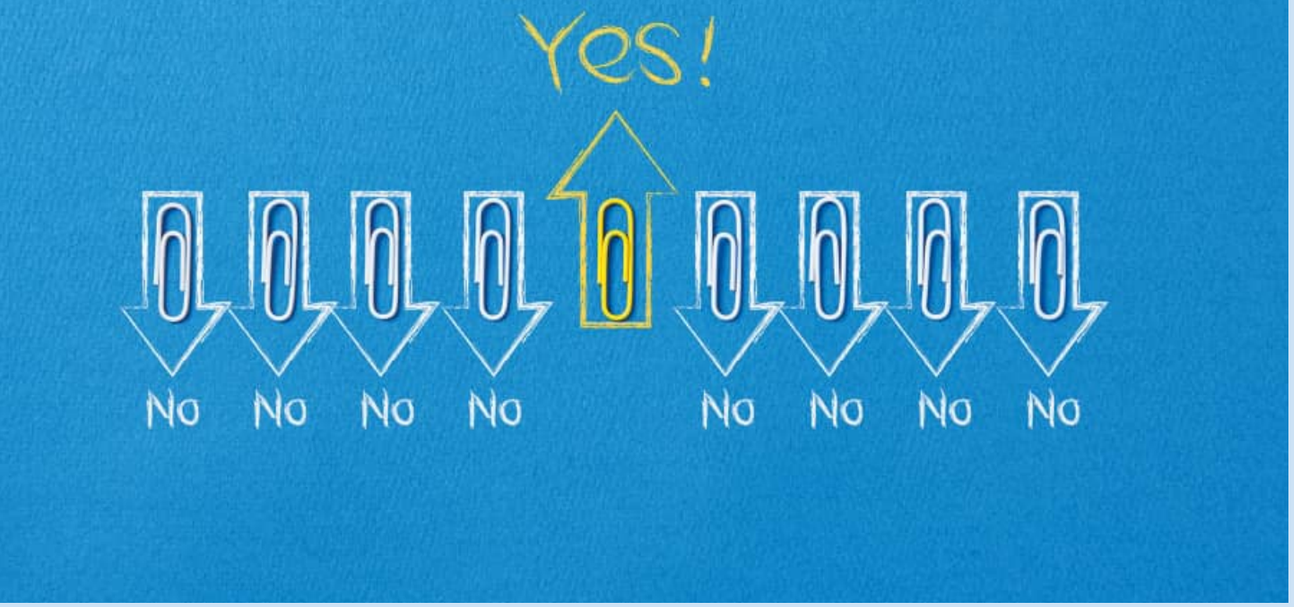
국어의 높임말 사용을 긍정적으로 보는 가장 큰 이유는 '웃어른을 공경하고 예의를 중시하는 우리 문화를 잘 담고 있음(41.8%)'이었으며, '사람들이 존경하는 마음과 공손함을 표시하는 중요한 수단임(27.0%)', '다양한 말투를 제공함으로써 한국어 표현력을 풍부하게 함(24.7%)'이 그 뒤를 이었습니다.

우리 국민의 46.9%가 욕설을 사용하고, 48.1%가 비속어를 사용하는 것으로 나타났습니다.

욕설이나 비속어를 사용하는 이유로 기분이 나쁠 때 그것을 표현하기 위해서(32.6%)라는 응답이 가장 많았고, 습관적으로(23.1%) 사용한다고 응답하였습니다.

데이터 앤 의사결정 리터러시 연결

강양석 | DEEP SKILL 대표



데이터 과학자도 의사결정 리터러시가 필요하다.

데이터는 오늘날 사회 전반에 걸쳐 필수적인 자산으로 자리 잡았다. 정부 정책, 기업 전략, 개인의 일상적 선택에 이르기까지 우리는 데이터를 기반으로 판단하고 행동하려는 경향이 점점 강화되고 있다. 그러나 데이터를 단순히 '읽을 수 있다'는 것만으로는 효과적인 의사결정을 내릴 수 없다. 아무리 정교한 분석이라도 그것이 무엇을 위해, 어떤 맥락에서, 어떻게 쓰일 것인지 명확하지 않으면 데이터는 쉽게 잘못된 확신으로 이어진다. 우리는 종종 데이터 리터러시를 갖춘 사람이 더 나은 결정을 내릴 것이라 기대하지만, 실제 현장에서는 데이터에 능숙한 사람이 오히려 판단 착오에 빠지기도 한다. 반대로, 복잡한 분석 기법은 몰라도 핵심 질문을 날카롭게 던지는 사람이 더 나은 전략을 제시하기도 한다. 그 차이는 의사결정 리터러시의 유무에서 비롯된다.

분석의 출발점은 데이터가 아니라 의사결정이어야 한다. 대부분의 분석은 데이터를 수집하고, 전처리하고, 다양한 통계기법을 적용한 후 시각화로 마무리된다. 이 과정이 잘 수행되었다고 해서 반드시 좋은 결정이 도출되는 것은 아니다. 오히려 잘못된 질문을 바탕으로 한 정교한 분석은 잘못된 결정을 정당화하는 도구가 되기도 한다. 한 국내 보험사는 고객 이탈률이 높아짐에 따라 빅데이터 기반 이탈 예측 모델을 구축했다. 고도화된 머신러닝 기법을 활용해 이탈 가능성이 높은 고객 리스트를 도출하고, 해당 고객들에게 할인 혜택과 맞춤형 메시지를 제공하는 마케팅 캠페인을 시행했다. 하지만 결과는 기대와 달리 이탈률 개선 효과는 미미했다. 원인은 간단했다. 이탈 예측은 성공했지만, 실제로 고객이 왜 이탈하는지를 의사결정 관점에서 해석하지 못했기 때문이다. 예를 들어 가격 외에 서비스의 복잡성, 고객센터의 응답 지연 등은 데이터에 포함되지 않은 정성적 요인이었고, 분석 결과는 그 맥락을 반영하지 못했다. 즉, 무엇을 분석할 것인가를 결정할 때, 의사결정자 관점에서 문제의 본질을 재정의하지 않고 데이터를 먼저 붙든 것이다.

반면, 미국 로스앤젤레스시는 범죄율을 줄이겠다는 단순 목표에서 출발하지 않았다. 대신 왜 특정 지역에서만 범죄가 집중되는가, 범죄 발생 전 조짐은 무엇인가라는 구체적 의사결정 질문을 던졌고, 그에 맞춰 데이터를 설계하고 수집했다. 경찰 순찰 기록, 조도 센서 데이터, 911 콜 기록, 빈집 정보 등 다양한 데이터를 통합해 범죄 발생 전후의 조건을 분석했고, 결과적으로 예방 순찰 중심의 전략으로 전환해 실제 범죄율을 낮추는 성과를 거뒀다. 이 사례는 분석이 아니라 의사결정의 재정의에서 출발했기에 데이터 분석이 효과적으로 작동한 예다.

데이터 과학자는 본래 기술 기반의 전문가로 인식되기 쉽다. 그러나 데이터 과학의 역할이 실제 세상의 문제를 해결하는 데 있다면, 기술적 능력만으로는 부족하다. 의사결정 리터러시란 데이터를 통해 무엇을 증명하고 어떤 선택을 해야 하는가를 구조적으로 사고하는 능력이다. 예컨대 기획자가 신제품을 출시할 시점을 결정해야 해요라고 말했을 때, 단순히 과거 3년간의 판매 데이터를 보여드릴게요라고 반응하는 분석가는 기술 중심이다. 반면 출시 시점이 중요한 이유가 수요 예측 때문인가요 아니면 경쟁사 움직임 때문인가요라고 되묻는 분석가는 의사결정 맥락을 이해하려는 사람이다. 후자의 분석가는 단지 데이터를 분석하는 것이 아니라 의사결정의 프레임을 함께 설계하는 파트너가 된다. 이처럼 데이터 과학자도 스스로가 다루는 분석이 의사결정과 어떤 관련을 갖는지 끊임없이 고민하고 있어야 한다.

다음 표는 기술 중심 데이터 과학자와 의사결정 리터러시를 갖춘 데이터 과학자의 차이를 비교한 것이다.

구 분	기술 중심 데이터 과학자	의사결정 리터러시를 갖춘 데이터 과학자
분석 시작점	분석 가능한 데이터 중심	해결해야 할 의사결정 중심
질문 방식	무엇을 분석할 수 있지	무엇을 알아야 결정할 수 있지
커뮤니케이션	지표 설명	판단 프레임 제안
성과 정의	예측 정확도	실제 영향력 실행 가능성

데이터 리터러시는 더 이상 분석가만의 기술이 아니다. 의사결정자 역시 데이터의 의미를 이해하고 분석 결과가 지닌 한계를 읽을 수 있어야 한다. 반대로 데이터 과학자도 자신이 생산하는 정보가 어떤 판단의 흐름에서 사용될지를 감각적으로 이해해야 한다. 이때 두 리터러시 사이의 연결은 일방향이 아닌 쌍방향 소통을 기반으로 한다. 즉 데이터를 해석하는 능력과 결정을 구조화하는 능력은 서로 다른 기술이 아니라 같은 사고 과정의 다른 면이다.

결론적으로 데이터 중심 사회에서 중요한 것은 데이터 그 자체보다 그것이 어떻게 판단과 연결되는가다. 분석 기술은 날로 발전하고 AI는 자동화를 앞당기고 있지만 여전히 무엇을 물어볼 것인가는 인간의 몫이다. 의사결정 리터러시는 데이터를 구조화된 질문의 언어로 바꾸고 불확실한 현실 속에서 판단 가능한 틀을 제공한다. 그리고 이 구조는 다음 장에서 다룰 논증 구조로 연결된다. 데이터와 판단 과학과 철학 분석과 해석의 교차점 우리는 이제 그 지점에서 사고해야 한다.

의사결정의 핵심은 논증구조이다.

의사결정이라는 것은 단순히 정보를 수집하고 그중 가장 타당한 선택지를 고르는 작업이 아니다. 실제 의사결정 과정은 복잡하고, 불완전한 정보 속에서 끊임없는 해석과 조정을 요구한다. 따라서 이를 효과적으로 수행하기 위해서는 사고의 구조를 명확히 이해하고, 다양한 주장과 관점을 조립하고, 조정하며, 최종적인 판단을 내리는 사고 프레임이 필요하다. 바로 이 지점에서 논증이라는 개념이 등장한다. 논증은 단순한 말싸움이나 의견 교환이 아니라, 주장을 설득력 있게 구성하고 타당하게 전개하는 일종의 사고의 건축물이라 할 수 있다. 그리고 이 논증을 통해 의사결정은 단순한 선택이 아닌 합리적인 판단의 결과로 자리 잡게 된다.

의사결정 리터러시의 핵심은 논증을 이해하고 다룰 수 있는 능력이다. 논증을 잘 구성한다는 것은 단순히 말을 논리적으로 잘하는 것을 넘어서, 주장의 기반이 되는 상황을 이해하고, 왜 그런 주장을 하는지 이유를 설명하며, 그 이유를 뒷받침하는 증거를 제시하고, 그 모든 것을 가능하게 만드는 전제까지 고려



하는 사고 구조를 갖춘다는 의미다. 논증의 구조는 보통 상황 상황에 대한 해석에서 출발하여 주장 주장을 통해 어떤 방향으로 가야 하는지를 제시하고, 그 주장을 뒷받침하는 이유 이유를 설명한 뒤, 그 이유를 지지하는 증거를 제시하고, 마지막으로 그 모든 것을 가능하게 하는 숨겨진 전제 전제로 구성된다.

이 다섯 가지 요소는 각각 독립적이면서도 유기적으로 연결되어 있다. 상황은 논의의 출발점이며, 주장은 그 상황에 대한 입장, 즉 결론이다. 이유는 주장을 뒷받침하는 논리적 근거이며, 증거는 그 이유를 정량적 정성적으로 지지하는 실질적 자료다. 전제는 논증이 작동하기 위한 조건 혹은 배경 지식으로, 보통은 말로 드러나지 않지만 핵심적이다. 다음은 각 요소의 기능과 중요도를 정리한 표이다.

논증 요소	기 능	설 명
상 황	문제 정의	어떤 맥락에서 의사결정이 필요한지 설명
주 장	방향 제시	무엇을 해야 하는가에 대한 결론
이 유	논리적 연결	주장이 왜 타당한지를 설명
증 거	타당성 확보	이유를 뒷받침하는 수치, 사례, 관찰 결과 등
전 제	숨겨진 전제	논리 구조가 작동하기 위한 기본 배경 지식, 관점

논증 구조에 대한 감각이 있는 사람과 그렇지 않은 사람의 주장은 명확히 구분된다. 예를 들어, “이 정책은 효과가 없다”는 주장을 할 때, 논증 구조를 이해한 사람은 “해당 정책이 효과가 없다는 주장은 최근 3년간 관련 수치가 하락세를 보이지 않았다는 증거에 근거한다. 이는 정책이 목표했던 성과 지표에 변화를 주지 못했다는 점에서 효과가 없다는 판단을 내리는 이유다. 이는 정책이 일정 수준 이상의 수치 개선을 동반해야 한다는 전제가 있다”라고 말한다. 반면 논증 구조를 잘 이해하지 못한 사람은 “다들 실패했다고 말하잖아요” 혹은 “그건 그냥 내 생각이야”처럼 개인적 판단이나 일반 여론에 의존하는 주장을 하게 된다.
조직 내 의사결정 과정에서도 이러한 차이는 분명하게 드러난다. 어떤 안건에 대해 찬반이 갈릴 때, 논증에 능숙한 사람은 자신의 입장을 명확히 표현하고, 타인의 주장에 대해 반론을 제기할 수 있으며, 새로운 대안을 제시하는 데 능하다. 즉, 논증은 단순히 정답을 찾아가는 직선적인 사고가 아니라, 주장을 내고, 그에 대한 반론을 받아들이며, 새로운 주장을 만들어내는 정반합의 순환적 사고 구조를 가능하게 한다. 이는 갈등을 회피하는 것이 아니라 갈등을 통해 더 나은 결론에 이르게 하는 길이다.

이러한 정반합적 사고는 특히 복잡하고 이해관계가 얽힌 조직의 의사결정 과정에서 더욱 중요하다. 예를 들어, 어느 도시가 신규 교통망 확충을 고민하는 경우, “혼잡하니 지하철 노선을 추가하자”는 주장과 “예산이 부족하니 민간 투자 우선이다”라는 반론이 부딪히게 된다. 이때, 논증 구조에 익숙한 의사결정자는 이 두 입장을 절충해 “예산 제약을 고려한 단계적 민관 협력 모델을 도입하되, 향후 교통량 증가 추세를 반영해 점진적 확장을 고려한다”는 식의 정합적 대안을 도출할 수 있다.
결국, 의사결정 리터러시란 복잡한 문제 속에서 다양한 관점과 주장을 구조적으로 통합해내는 능력이며,

논증 구조에 대한 감각은 그 핵심이다. 우리는 흔히 데이터를 중심으로 한 '증거'를 강조하지만, 그 증거가 어디에서 나왔고 무엇을 지지하며 어떤 논리를 따라가는지 이해하지 못하면 오히려 판단을 왜곡시킬 수도 있다. 그래서 다음 장에서는 이 논증 구조 안에서 데이터 리터러시가 어떤 역할을 수행하는지를 본격적으로 살펴볼 것이다. 데이터는 단지 사실의 나열이 아니라, 주장을 강화하고 의사결정을 정교하게 만드는 중요한 사고의 도구이기 때문이다.

논증구조 관점으로 데이터 분석을 이해하자.

데이터 리터러시를 이야기할 때 우리는 주로 데이터를 정확히 읽고 해석하고 시각화할 수 있는 기술을 떠올린다. 하지만 그것은 리터러시의 절반일 뿐이다. 데이터를 '잘 읽는 것'보다 더 중요한 것은 그 데이터가 어떤 주장과 판단 구조 안에서 작동하는지를 이해하는 것이다. 다시 말해 데이터 리터러시는 고립된 기술이 아니라 논증이라는 사고 틀 안에서 재해석되어야 한다. 앞서 살펴본 논증 구조, 즉 상황 주장 이유 증거 전제라는 구성 안에서 데이터를 바라볼 때 비로소 데이터의 역할이 명확해지고, 데이터 분석 결과가 설득력 있는 의사결정으로 연결될 수 있다.

일반적으로 데이터 분석은 논증 구조 중 '증거와 이유'의 사이를 메우는 역할을 수행한다. 예를 들어 "재택근무가 생산성을 높인다"는 주장에 대해 "업무 몰입도가 향상되기 때문이다"라는 이유가 주어졌다면, 그 이유를 지지하기 위한 증거로 "지난 6개월간 재택근무자의 프로젝트 완료율이 평균 대비 12% 높았다"는 데이터를 제시할 수 있다. 이때 데이터는 그저 사실의 나열이 아니라 이유를 지지하는 '정당성의 도구'로 쓰인다. 즉 데이터는 논증의 가장 설득력 있는 지점인 '이유의 증거화'를 통해 논리를 강화한다. 여기서 중요한 점은 데이터가 단순히 주장 자체를 정당화하는 것이 아니라, 주장과 이유 사이의 연결을 강화하는 방식으로 작동한다는 것이다.

하지만 데이터의 역할은 단지 '증거'에 국한되지 않는다. 논증의 모든 요소는 데이터 과학과 연결될 수 있으며, 각 요소마다 데이터 과학자가 기여할 수 있는 방식이 다르다. 예를 들어 상황 정의에서 데이터는 탐색적 분석과 외부 환경 분석을 통해 문제의 실체를 규명하는 데 도움을 줄 수 있다. 주장의 단계에서는 시나리오별 예측 모델링을 통해 선택지를 비교할 수 있다. 이유의 단계에서는 변수 간 인과관계를 분석하거나 가설 검정을 통해 논리적 연결을 점검할 수 있다. 그리고 증거는 말할 것도 없이 실증적 자료를 제공하는 주된 수단이다. 마지막으로 전제는 문화적 맥락이나 조직의 철학처럼 수치화되기 어려운 요소이지만, 텍스트 마이닝이나 정성 데이터 분석을 통해 일부 드러낼 수 있는 가능성도 있다.

아래는 논증 요소별로 데이터 과학이 기여할 수 있는 방식과, 각 요소에서 분석 시 고려해야 할 주요 사항을 정리한 표이다.

논증 요소	데이터 과학적 기여	설 명
상 황	환경 스캐닝, 탐색적 데이터 분석	문제 정의의 적절성, 맥락 반영 여부
주 장	시나리오 모델링, 대안 비교	정책/전략의 비교 기준, KPI 설정
이 유	인과분석, 변수 간 관계 분석	설명력, 논리적 일관성
증 거	실측 데이터, 통계 지표	신뢰도, 대표성, 오류 가능성
전 제	문화 분석, 텍스트 기반 맥락 분석	명시되지 않은 가정, 조직적 이해관계

이처럼 논증 요소 하나하나가 데이터 과학적 관점과 연결될 수 있으며, 분석가는 자신의 작업이 이 중 어디에 위치하는지를 명확히 인식해야 한다. 자신이 수행한 분석이 단순히 데이터를 정리한 것인지, 아니면 어떤 이유를 설명하는 것인지, 혹은 전제를 흔드는 해석을 제공하는 것인지 스스로 질문하는 과정이 필요하다.

예를 들어 한 교육 기관이 "온라인 강의가 오프라인 강의보다 효과적이다"라는 주장을 펴고자 할 때, 단순히 만족도 조사를 그래프로 제시하는 것만으로는 부족하다. 이때 논증 구조 안에서 데이터를 배치하면 전체 흐름이 달라진다. 상황은 "코로나19 이후 온라인 강의가 일상화되었다"이며, 주장은 "온라인 강의가 학습 효과 측면에서 더 낫다"이다. 이유는 "개별 진도 조절과 반복 학습이 가능하기 때문"이고, 증거는 "반복 수강 비율이 높은 학습자의 평가 점수가 평균보다 15% 높았다"라는 분석 결과다. 전제는 "학습 효과는 학습자 만족도보다 성취도로 평가해야 한다"는 암묵적 기준이다. 이 논증 안에서 데이터 분석은 단순한 시각화가 아니라, 이유와 전제에 논리적 설득력을 부여하는 핵심 역할을 하게 된다.

결국 데이터 리터러시는 단순한 통계 기법 습득을 넘어, 데이터를 의사결정의 구조 속에서 사고하는 능력으로 확장되어야 한다. 데이터는 그 자체로 아무것도 주장하지 않는다. 주장을 설계하고 구조화하는 것은 사람의 몫이며, 데이터는 그 논리적 구조 속에서 설득력을 더하는 보강재일 뿐이다. 따라서 우리는 이제 데이터 리터러시를 '논증 구조 속 사고의 언어'로 다시 정의해야 한다. 그래야만 데이터 과학은 단순한 도구가 아니라 판단의 체계로 기능할 수 있다. 다음 장에서는 이러한 논증 구조와 더욱 직접적으로 연결된 추론 방식과 데이터 기법들의 관계를 살펴볼 것이다. 추론은 논증의 핵심 작동 방식이며, 데이터 분석 기법 역시 다양한 추론 구조와 닮아 있기 때문이다.





분석도 추론이다.

논증이 설득력 있게 작동하기 위해서는 단지 주장을 나열하고 근거를 덧붙이는 것을 넘어서, 그 주장과 근거 사이를 연결하는 추론의 과정이 명확해야 한다. 추론은 주어진 정보로부터 새로운 결론을 이끌어내는 사고 방식으로, 논증의 핵심 작동 메커니즘이라 할 수 있다. 우리는 일상적으로 추론을 사용해 판단을 내린다. 그러나 이 추론이 어떤 방식으로 작동하는지 자각하지 않으면, 논증은 쉽게 비약으로 흐르거나, 오류가 포함된 채 받아들여질 수 있다. 데이터 분석 역시 마찬가지다. 데이터를 분석하는 대부분의 기법은 단순한 수치 계산이 아니라, 일정한 전제에서 출발해 의미 있는 결론을 이끌어내려는 추론의 도구다. 따라서 데이터 분석 기법을 제대로 이해하기 위해서는 그것이 어떤 추론 구조를 따르고 있는지를 파악할 필요가 있다.

추론은 크게 네 가지 방식으로 구분할 수 있다. 첫째는 연역적 추론으로, 일반적인 원리나 법칙에서 개별 사례를 이끌어내는 방식이다. 예를 들어 모든 사람은 죽는다, 소크라테스는 사람이다, 따라서 소크라테스는 죽는다 같은 구조가 이에 해당한다. 둘째는 귀납적 추론으로, 여러 개별 사례로부터 일반적인 결론을 도출하는 방식이다. 세 번째는 가추적 추론인데, 이는 관찰된 현상에 가장 잘 들어맞는 설명을 역으로 구성하는 방식이다. 즉 가장 그럴듯한 원인을 추정해내는 것이다. 마지막으로 유추적 추론은 유사한 구조나 성질을 갖는 두 대상을 비교해, 한 대상의 특성을 다른 대상에도 적용하는 방식이다. 예를 들어 A학교에서 시도한 교육 정책이 성공했다면, 유사한 조건을 가진 B학교에서도 효과가 있을 것이라는 추론이 여기에 해당한다.

이러한 추론 방식은 데이터 분석 기법의 작동 원리와 매우 밀접하게 닮아 있다. 예를 들어 회귀 분석은 특정 변수들이 결과 변수에 어떤 영향을 미치는지를 추정할 때, 기본적으로 인과적 연역 또는 귀납적 구조를 따르고 있다. 시계열 분석은 시간 흐름에 따른 패턴을 일반화하려는 귀납적 추론의 대표적인 예다. 클러스터링은 가추적 사고와 유사하다. 주어진 데이터 속에서 어떤 패턴이나 그룹이 존재하는지, 즉 관찰된 특성을 설명할 수 있는 잠재적 구조를 역으로 구성해내기 때문이다. 분류(Classification) 기법은 유추적 추론과 닮아 있다. 이미 분류된 사례들과 유사한 특성을 가진 새로운 사례가 어떤 그룹에 속할지를 판단하는 과정이기 때문이다. 마지막으로 인과 추론은 명시적 전제와 개입 가능성을 고려하여 가장 논리적인 결론을 이끌어내는 연역적 추론의 전형에 가깝다.
다음 표는 추론 방식과 주요 데이터 분석 기법 간의 관계를 정리한 것이다.

추론 방식	개 념	유사한 데이터 분석 기법	설 명
연역	일반 → 개별	인과 추론, 가설 검정	전제를 기반으로 특정 결론을 도출
귀납	개별 → 일반	시계열 분석, 통계 요약	반복된 관측을 바탕으로 일반화
가추	가장 그럴듯한 설명	클러스터링. 이상탐지	관찰된 현상에 맞는 설명 구성
유추	유사성 기반 일반화	분류(classification), 추천 알고리즘	유사한 사례의 속성을 확장 적용

이처럼 데이터 분석 기법은 단지 기술적 도구가 아니라, 각각의 추론 방식을 체계화한 사고 방식의 표현이다. 그래서 분석가는 단지 어떤 기법을 쓸 것인지가 아니라, 그 기법이 어떤 추론을 내포하고 있는지를 이해해야 한다. 그래야만 분석의 설계가 의사결정의 논증 구조와 자연스럽게 연결되고, 분석 결과 역시 판단 가능한 형태로 전달될 수 있다.

예를 들어 어떤 지방정부가 “청년 창업 지원이 지역 경제를 활성화시킨다”는 주장을 평가하고자 할 때, 다양한 추론 방식이 복합적으로 활용된다. 연역적 추론은 기존 문헌에 따라 창업 증가 → 고용 증가 → 경제 활성화라는 인과 체계를 전제로 정책 효과를 측정하려 한다. 귀납적 추론은 과거 비슷한 정책 시행 이후의 지역 경제 지표 변화를 살펴본다. 가추적 추론은 지역별 산업 생태계나 창업 생존율과 같은 특수 조건을 분석해 효과가 있는 지역과 없는 지역의 차이를 설명하려 한다. 유추적 추론은 유사한 경제 규모나 인구 구조를 가진 타 지자체의 사례를 근거로 효과를 예상한다. 이처럼 실제 의사결정 과정에서는 다양한 추론이 혼합적으로 작동하며, 각각은 다른 데이터 기법과 연계된다.

결국 데이터 분석을 잘한다는 것은 단지 정확하게 계산하고 깔끔하게 시각화하는 것을 의미하지 않는다. 그것은 논증 구조 안에서 자신의 분석이 어떤 추론을 기반으로 작동하고 있는지를 이해하고, 그 추론의 타당성과 한계를 스스로 점검할 수 있는 사고력을 갖추는 것을 의미한다. 분석가는 단지 데이터를 처리하는 엔지니어가 아니라, 복잡한 판단의 맥락에서 어떤 추론이 합리적인지를 설계하고 검증하는 사고 설계자여야 한다. 다음 장에서는 이러한 관점을 바탕으로, 데이터 과학자가 더 높은 수준의 의사결정 능력을 갖기 위해 갖추어야 할 태도와 시대적 배경, 그리고 왜 지금 두 리터러시의 통합적 식견이 절실한지 정리할 것이다.

데이터 과학자는 어떻게 의사결정 리터리시를 효과적으로 가질까?

지금 이 시대에 데이터 과학자에게 요구되는 역량은 이전과는 본질적으로 다르다. 과거에는 통계 기법에 능하고, 프로그래밍을 잘 다루며, 분석 도구를 다룰 수 있는 사람이면 충분했다. 그러나 이제는 단순히 데이터를 처리하는 능력만으로는 중요한 역할을 할 수 없다. 데이터 분석 그 자체보다 중요한 것은 그것이 '무엇을 위한 것인가'를 이해하고, 그 목적을 위해 사고를 구조화하고 판단할 수 있는 능력이다. 다시 말해, 데이터 과학자에게는 더 이상 데이터 리터러시만으로 충분하지 않으며, 의사결정 리터러시를 함께 갖춘 통합적 사고력이 필수적인 시대가 된 것이다.

이러한 변화의 배경에는 기술의 진화가 자리 잡고 있다. 특히 인공지능 기술의 급속한 발전은 전통적인 분석 기술의 가치에 근본적인 의문을 던지고 있다. 예전에는 분석가가 스스로 모델을 설계하고 데이터를 정제하며 가설을 검증해야 했지만, 이제는 많은 부분을 AI가 자동화할 수 있다. 예측 모델, 분류 모델, 군집화 분석 같은 기법은 대부분의 경우 오픈소스 툴과 자동화된 AI 플랫폼에서 기본 제공된다. 분석을 잘하는 사람과 그렇지 않은 사람 사이의 기술 격차는 빠르게 줄어들고 있다. 반면, 어떤 질문을 던질 것인가, 어떤 논리 구조로 판단을 설계할 것인가, 그리고 어떤 이해관계 속에서 결정을 도출할 것인가라는 문제는 여전히 인간의 사고력에 달려 있다.

그래서 최근 데이터 과학의 중심이 점점 소프트 스킬 중심으로 이동하고 있다. 단지 코드를 잘 짜는 능력보다, 복잡한 문제를 구조화하고, 팀 내에서 설득력 있게 의견을 제시하며, 타인의 관점을 이해하고 조율할 수 있는 능력이 더 높은 평가를 받는다. 실제로 많은 조직에서 데이터 분석가 채용 시, 기술 테스트뿐 아니라 커뮤니케이션 능력, 프레젠테이션 역량, 맥락 이해도 등을 함께 평가한다. 특히 기업의 전략 기획 부서나 정책 결정과 연계된 공공기관에서는 분석 역량보다 문제 정의 능력과 판단 구조 설계 능력을 더 중요하게 본다. 이 변화는 단지 기술의 변화가 아니라 직무의 본질적 변화를 의미한다.





이러한 배경에서 데이터 과학자가 더 높은 의사결정 능력을 갖추기 위해서는 다음과 같은 사고 전환이 필요하다. 첫째, 데이터를 중심으로 사고하는 것이 아니라, 결정을 중심으로 사고하는 태도다. 분석은 판단을 위한 도구이지 목표가 아니다. 둘째, 데이터 분석을 논증 구조 안에 배치할 수 있어야 한다. 상황 주장 이유 증거 전제라는 논증 프레임 속에서 내가 생산한 정보가 어떤 역할을 하는지를 자각해야 한다. 셋째, 다양한 추론 구조에 대한 이해를 바탕으로 내가 사용하고 있는 분석 기법이 어떤 사고 메커니즘을 반영하고 있는지를 해석할 수 있어야 한다. 넷째, 의사결정자와의 대화를 통해 기술 언어를 전략 언어로 번역할 수 있어야 한다. 분석 결과를 잘 설명하는 것이 아니라, 분석의 핵심이 의사결정에 어떤 방식으로 기여할 수 있는지를 명확히 전달해야 한다.

이 모든 것은 결국 리터러시의 통합이라는 문제로 귀결된다. 데이터 리터러시와 의사결정 리터러시는 분리된 능력이 아니다. 앞서 보았듯 데이터는 이유와 증거를 강화하지만, 전제와 주장, 상황을 설계하는 것은 사고력의 몫이다. 이 둘을 함께 다룰 수 있을 때, 분석은 단지 정확한 결과의 제시가 아니라, 설득 가능한 판단의 구조를 만드는 작업으로 전환된다. 그리고 이것이야말로 오늘날 데이터 과학자가 의사결정자로 성장하기 위해 반드시 갖추어야 할 역량이다.

우리는 데이터 중심 사회를 살아가고 있다. 그러나 동시에 판단의 책임은 여전히 인간에게 남아 있다. AI는 무한한 계산 능력을 제공하지만, 무엇을 계산할 것인가는 결국 사람의 몫이다. 데이터를 다루는 모든 사람은 이제 '정보를 제공하는 사람'이 아니라, '판단을 설계하는 사람'이 되어야 한다. 그 시작은 바로 데이터 리터러시와 의사결정 리터러시를 하나의 사고 체계로 통합해 이해하는 데 있다. 그것이 이 시대의 데이터 과학자에게 요구되는 가장 중요한 식견이다.



AI를 활용하여 서베이 및 고객 데이터 분석하기

구자룡 | 밸류바인 대표

빅데이터와 인공지능 기술의 발전으로 시장 및 고객 분석 방법이 혁신적으로 변화하고 있다. 특히 서베이 데이터와 AI의 결합은 시장의 특성을 더 정확하고 깊이 있게 이해할 수 있는 새로운 기회를 제공한다. 또한 고객 데이터와 AI의 결합은 머신러닝과 모델링에 대한 이해가 낮은 실무자들도 군집화 분석을 통해 타깃 마케팅을 가능하게 한다. 이 글에서는 AI를 활용하여 서베이 데이터를 수집하고 분석하는 방법과 이를 통해 시장 특성을 파악하고 의사결정에 활용하는 프로세스를 살펴본다. 그리고 고객 데이터로 군집화 분석을 통해 타깃팅 하는 방법에 대해 설명한다. 기업이 빅데이터 시대에 경쟁 우위를 확보하기 위해서는 전통적인 데이터 분석 방법론에 AI 기술을 접목하여 보다 스마트한 의사결정을 내리는 것이 필수적이다.

빅데이터 시대에 시장 조사와 서베이를 하는 이유

디지털 혁신으로 인해 우리는 매일 2.5퀸틸리언 바이트(약 2,500조 테라바이트는 약 500경(京) 장의 사진을 저장할 수 있는 용량)의 데이터가 생성되는 빅데이터 시대에 살고 있다. 소셜 미디어 포스트, 온라인 구매 기록, IoT 기기의 센서 데이터, 클릭 로그 등 기업들은 이전에 없던 방대한 양의 고객 데이터에 접근할 수 있게 되었다. 이런 환경에서 일부는 "전통적인 시장 조사와 서베이가 여전히 필요한가?"라는 의문을 제기한다. 답은 단호하게 "그렇다" 이다.

빅데이터는 '무엇이' 일어났는지를 보여주지만, '왜' 일어났는지를 설명하는 데는 한계가 있다. 예를 들어, 온라인 쇼핑 데이터는 고객이 어떤 제품을 클릭하고 구매했는지 알려주지만, 왜 그 제품을 선택했는지, 구매 과정에서 어떤 고민을 했는지, 만족도는 어떤지 등의 맥락 정보는 제공하지 않는다. 이 맥락과 동기를 이해하기 위해서는 직접적인 질문과 응답이 필요하다.

또한 서베이는 빅데이터의 사각지대를 보완한다. 디지털 행동 데이터는 온라인 활동을 하는 사람들에 국한되어 있으며, 특정 인구 집단(예: 디지털 소외 계층)을 과소 대표할 수 있다. 반면 잘 설계된 시장 조사는 대표성을 확보하여 더 포괄적인 인사이트를 제공한다.

시장 조사의 또 다른 중요한 역할은 가설 검정이다. 기업은 고객이나 시장에 대한 다양한 가설을 가지고 있지만, 이를 체계적으로 검정하지 않으면 편향된 의사결정을 내릴 위험이 있다. 서베이는 이러한 가설을 객관적으로 검정할 수 있는 도구를 제공한다.

불확실성이 높은 현대 비즈니스 환경에서 데이터에 기반을 둔 의사결정은 더 이상 선택이 아닌 필수가 되었다. 맥킨지의 연구에 따르면 데이터 기반 의사결정을 하는 기업은 그렇지 않은 기업보다 23% 더 높은 성과를 달성한다고 한다. 서베이 데이터는 이러한 데이터 기반 의사결정의 핵심 요소 중 하나이다.

고객 경험이 경쟁 우위의 핵심 요소로 부상하면서, 고객의 목소리를 직접 듣는 것의 중요성이 더욱 커지고 있다. 고객만족도(CSI, Customer Satisfaction Index), 고객노력지수(CES, Customer Effort Score), 고객순추천지수(NPS, Net Promoter Score) 등의 지표는 서베이를 통해 측정되며, 이는 고객 경험 관리의 기초가 된다.

결국 서베이는 '숫자 너머의 이야기'를 들을 수 있게 해주며, 이는 오늘날처럼 복잡하고 빠르게 변화하는 시장에서 전략적 판단을 내리는 데 필수적인 도구로 자리 잡고 있다.



AI를 활용하여 시장 조사와 서베이 데이터 수집하기

과거에는 서베이를 설계하고 실행하는 일이 전문적인 조사 기관의 몫이었다. 조사 설계자는 수많은 참고 문헌과 과거 데이터를 분석하여 문항을 구성하고, 그 문항들을 실제 고객에게 전달하기 위한 플랫폼을 개발하거나 외부 패널을 모집해야 했다. 설문지 완성부터 데이터 수집까지 몇 주에서 몇 달이 걸리는 경우도 흔했다. 하지만 AI의 등장으로 이 모든 과정이 훨씬 빠르고 간편해졌다.

생성형 AI, 특히 챗GPT와 같은 도구를 활용하면 누구나 조사 목적에 맞춘 문항을 손쉽게 생성할 수 있다. 예를 들어 "20대 직장인을 대상으로 자사 커피 브랜드에 대한 인식을 파악하고 싶다"는 조사 목적을 입력하면, AI는 관련 배경 질문(연령, 직업 등)부터, 인식 측정 문항(브랜드 친숙도, 신뢰도, 재구매 의향 등)까지 자동으로 구성해준다. 그뿐만 아니라 응답 방식도 객관식, 리커트 척도, 서술형 중 적절한 형식을 제안하며, 설문 흐름도 자연스럽게 이어지도록 도와준다.

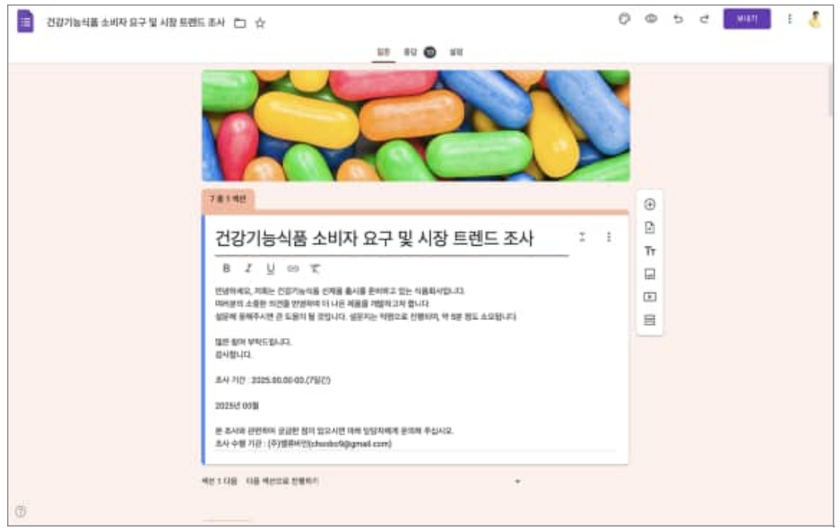
AI는 설문 플랫폼과도 손쉽게 연동된다. 구글 설문지(Google Forms), 네이버 폼, 카카오 폼과 같은 설문 툴에 AI가 생성한 문항을 바로 적용하거나, 붙여넣기를 할 수 있으며, API를 통해 설문 배포와 응답 수집을 자동화할 수도 있다. 특히 고객관계관리(CRM) 시스템과 연동하면, 특정 조건(예: 최근 30일 내 구매 고객, VIP 고객 등)에 맞는 응답자만을 자동으로 설정할 수 있다. 이는 응답률과 응답 신뢰도를 동시에 향상시킨다.

예를 들어, 건강기능식품에 대한 신제품 개발을 목적으로 시장조사를 하고자 할 때 챗GPT의 프롬프트에 “건강기능식품 시장에서 소비자 요구, 경쟁 상황, 시장 트렌드를 파악하여 신제품 출시 전략을 수립하고자 한다.”라는 목표를 제시하고 주요 질문으로 “소비자들이 건강기능식품에서 원하는 것은 무엇인가? 주요 경쟁 제품과의 차별화 요소는 무엇인가? 적정 가격대 혹은 가격 감수성 테스트로 가격대는 어떠한가? 현재 시장 트렌드는 어떠한가?” 등을 요청할 수 있다. 그리고 조사 방법은 “구글설문지를 이용하여 온라인 조사”라고 구체적으로 작성한 다음 “이런 내용으로 설문 조사를 하려고 해. 구글설문지 폼에 맞게 설문지를 만들어줘”라고만 하면 어느 정도 활용



[그림1] 챗GPT를 이용한 질문지 설계

가능한 질문지를 구성해준다[그림1]. 이를 바탕으로 연구자 혹은 마케터는 구글 설문지에 가서 챗GPT가 제안한 내용을 참고하여 조사 목적에 부합하는 구체적이고 체계적인 질문을 작성하면 된다[그림2].
최근에는 인터랙티브 챗봇 방식의 서베이도 각광받고 있다. 이는 기존의 고정된 설문지 방식이 아닌, AI가 응답자의 답변에 따라 후속 질문을 유동적으로 구성하는 방식이다. 예를 들어 응답자가 "제품 포장에 불만이 있다"고 답하면, AI는 "어떤 점이 불편했는지"를 추가로 질문하고, 이를 텍스트 응답으로 수집하여 자동 분류·요약할 수 있다. 이러한 방식은 특히 UX 조사, 서비스 개선을 위한 VOC(Voice of Customer) 분석 등에 효과적이다.

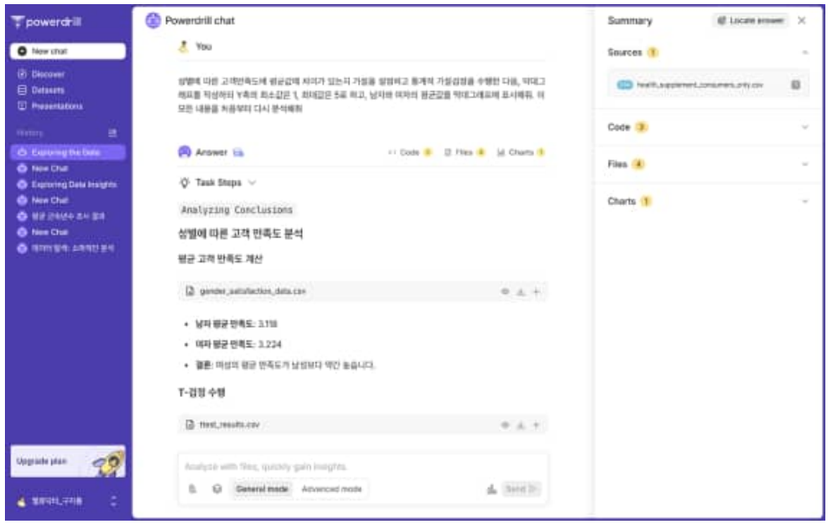


[그림2] 구글 설문지를 이용한 서베이 조사

AI를 활용하여 시장의 특성 분석하기

수집된 서베이 데이터는 그 자체로는 단순한 숫자와 텍스트의 모음에 불과하다. 이를 의미 있는 인사이트로 변환하기 위해서는 체계적인 분석이 필요하다. 과거에는 분석을 위해 전문 통계 소프트웨어(SPSS, SAS 등)와 통계 지식이 필수였지만, AI의 도움을 받으면 누구나 쉽게 데이터를 다룰 수 있다. 오픈AI의 챗GPT, 구글의 제미나이(Gemini), 마이크로소프트의 코파일럿(Copilot), 그리고 데이터 분석 전문 AI인 줄리어스(Julius.ai)와 파워드릴(Powerdrill.ai) 등은 자연어로 분석 지시를 내리면 자동으로 통계 처리와 시각화까지 수행해 준다.

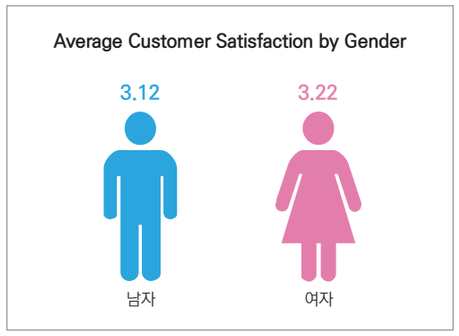
AI는 전통적인 통계 분석을 넘어 복잡한 패턴을 발견하고 다차원적인 시장 특성을 파악하는 데 탁월한 능력을 보여준다. 예를 들어, 고객 만족도 조사 데이터를 업로드한 후 "문항별 평균과 표준편차를 구해줘"



[그림3] 파워드릴 AI를 이용한 통계적 가설 검정 분석

(데이터 파일을 파워드릴에 업로드 후 프롬프트에 "성별에 따른 고객만족도 차이에 대한 통계적 가설 검정을 수행해줘"라고 요청)

라고 지시하면, AI는 즉시 기술통계 분석을 수행하고 표 형태로 결과를 정리한다. 이후 "성별에 따른 고객만족도의 평균값에 차이가 있는지 분석해줘"라고 하면, 자동으로 t-검정 분석을 수행하고 유의확률(p-value)도 함께 제공한다[그림3][그림4]. 이는 분석가가 데이터 분석의 기초 원리를 이해하고 있다는 전제 하에, 훨씬 빠르고 효율적인 분석이 가능함을 의미한다.



[그림4] 파워드릴 AI를 이용한 통계적 가설 검정 분석 결과

(남자 평균 만족도: 3.12, 여자 평균 만족도: 3.22로 여성의 평균 만족도가 남성보다 약간 높다. T-검정 결과 t-통계량: -0.956, p-값: 0.340으로 p-값이 0.05보다 크므로, 성별에 따른 고객 만족도 차이는 통계적으로 유의미하지 않다.)

또한 다중 회귀분석을 통해 만족도나 재구매 의향에 영향을 미치는 요인을 도출할 수 있으며, 변수의 중요도(표준화 회귀계수)까지 시각적으로 제시해준다. 이는 마케팅 전략 수립 시 '무엇에 집중해야 하는가?'에 대한 명확한 답을 제공한다. 예를 들어, 가격 만족도보다는 직원 응대가 재구매 의향에 더 큰 영향을 준다면, 내부 교육이나 서비스 매뉴얼 개선이 우선 전략이 될 수 있다.

AI를 활용하여 고객 군집화 분석으로 타깃팅하기

고객이 어떤 제품을 클릭하고 구매했는지 알려주는 고객의 행동데이터는 시장이 아니라 기업 내부에 있다. 서베이 데이터만으로는 한계가 있기 때문에 디지털 고객 데이터가 필요하다. 대부분 고객 데이터는 빅데이터다. 고객 데이터로 고객을 군집화하면 보다 정밀한 타깃 마케팅 전략을 수립할 수 있다. AI는 이러한 군집화 작업에서도 매우 유용하다. 군집화란 유사한 특성을 지닌 응답자들을 그룹으로 묶는 분석 기법으로, 대표적으로 K-평균 군집화(K-means Clustering) 분석, 계층적 군집화(Hierarchical Clustering) 분석 등의 알고리즘이 사용된다.

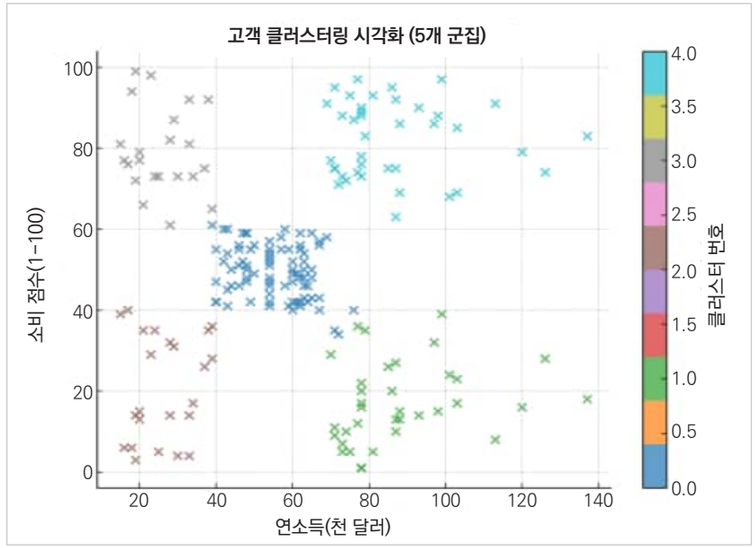
군집별 특성 분석을 통해 마케팅 타깃을 명확히 설정할 수 있다. 분석 결과, 1군집은 '합리적 소비형', 2군집은 '감성 충성형', 3군집은 '신규 체험형'이라면, 각각의 그룹에 맞는 메시지와 제품 패키징, 광고 채널을 다르게 설정할 수 있다. AI는 각 군집에 어떤 마케팅 전략이 효과적인지도 시뮬레이션 분석을 통해 예측할 수 있다.

예를 들어, 온라인 쇼핑몰의 고객 데이터베이스에서 고객ID, 성별, 연령, 연소득, 소비점수 등의 필드(변수)를 추출한다. 이 데이터를 기반으로 AI는 고객들을 3~5개의 군집으로 분류한다. 각 군집의 특성



[그림5] 챗GPT를 이용한 K-평균 군집화 분석

(쇼핑몰 고객 데이터를 챗GPT Data Analyst에 업로드 한 후 프롬프트에 "아래 순서대로 군집화 분석을 단계적으로 수행해줘"라고 요청)



[그림6] 챗GPT를 이용한 K-평균 군집화 분석 결과

(5개의 고객 세그먼트로 나누는 것이 적절하며, 각 고객 그룹이 소득 수준과 소비 성향에 따라 뚜렷하게 분리된 것을 확인할 수 있음)

(평균값, 표준편차 등)은 자동으로 프로파일링 되며, 시각화된 산점도나 클러스터 맵으로 분석 결과를 직관적으로 확인할 수 있다. 그림과 같이 쇼핑몰 고객 데이터로 K-평균 군집화 분석을 챗GPT로 수행했을 때 5개의 군집으로 세분화할 수 있었다[그림5][그림6].

군집화 분석 결과를 정리하면, '클러스터 0'은 중간 소득의 중장년층으로 이 군집에는 일반형 캠페인, 신규 고객 대상 유지 전략 등으로 전략 제안을 할 수 있다. '클러스터 1'은 고소득 & 고소비의 젊은 층으로 이 군집에는 VIP 혜택, 프리미엄 서비스 등으로 전략 제안을 할 수 있다.

예를 들어, 중간 소득의 중장년층(클러스터 0)을 대상으로 일반형 캠페인을 기획한다면, "조금 더 나은 하루를 위한 제안"이란 이름으로 월별 정기 프로모션(브랜드별 콜라보)을 생각해 볼 수 있다. 제공 혜택은 생활 밀착형 상품 최대 20% 할인 + 포인트 적립 2배, 채널은 문자, 이메일, 오프라인 지점 POP 등, 특징은 중장년층이 선호하는 품목 중심(건강, 리빙, 식품)으로 "고객 맞춤 추천 상품" 리스트를 제공하고, 구매 금액에 따라 고급 기프트(예: 건강식품 샘플, 머그잔 세트 등)를 제공하는 것이다. 이 캠페인의 기대 효과는 구매 빈도 증가와 브랜드 신뢰도 강화 등이 될 것이다.

군집화는 단순히 고객을 나누는 데 그치지 않는다. 신제품 타깃팅, 캠페인 반응 예측, 가격 민감도 분석 등에서도 중요한 기준이 된다. AI는 이러한 분석을 반복적으로 수행하며 지속적으로 세그먼트를 고도화

할 수 있도록 도와준다. 특히 마케팅 자동화 플랫폼과 연동하면 각 군집에 맞는 콘텐츠를 자동으로 제공하는 '맞춤형 마케팅'도 가능해진다.

데이터 기반 의사결정을 위한 AI 데이터 분석하기

서베이 데이터와 고객 데이터는 궁극적으로 기업의 의사결정에 반영되어야 한다. 단순히 보고서에 머무는 것이 아니라, 실질적인 전략 수립, 실행 계획, 자원 배분, 성과 평가에까지 연결되어야 그 가치를 발휘한다. AI는 바로 이 의사결정 과정을 보다 정밀하고 예측 가능하게 만들어준다.

예를 들어, 신제품 출시 전 콘셉트 테스트에서 얻은 데이터를 AI가 분석한 결과, A형 콘셉트는 20대 여성 고객에게, B형 콘셉트는 30대 직장인 남성에게 호응이 높다는 결과가 나왔다면, 타깃군을 구분하여 두 가지 버전으로 마케팅을 진행할 수 있다. 또한 A/B 테스트의 결과를 기반으로 광고 소재나 메시지를 조정하고, 캠페인 후 응답 결과를 실시간으로 분석하여 광고 집행 전략을 최적화할 수 있다.

무엇보다 중요한 것은, AI는 인간의 판단을 대체하는 것이 아니라 보완하고 확장하는 도구라는 점이다. AI는 빠르고 정확하게 정보를 제공하지만, 그 정보를 바탕으로 어떤 결정을 내릴지는 여전히 인간의 몫이다. 따라서 분석가는 AI와의 협업을 통해 더 넓고 깊은 관점에서 문제를 바라보고, 더욱 창의적이고 전략적인 해법을 도출해낼 수 있어야 한다.

결론적으로, AI와 데이터 분석의 결합은 기업이 고객을 더 잘 이해하고, 시장의 흐름을 더 정확히 파악하며, 전략적 의사결정을 더욱 효과적으로 내릴 수 있도록 돕는다. 데이터는 많아졌지만, 통찰은 여전히 부족한 시대다. AI는 그 간극을 메워줄 수 있는 강력한 도구다. 그러나 그 도구를 어떻게 활용하느냐는 전적으로 우리의 몫이다.



참고문헌

구자룡(2025).『챗GPT로 시작하는 데이터 리터러시』. 마들렌북.
구자룡(2024).『AI 데이터 분석』. 커뮤니케이션북스.
구자룡(2024).『데이터 마인드 기르는 습관』. 좋은습관연구소.
이명식·구자룡·양석준(2017).『마케팅 리서치(개정판)』. 형설출판사.
McKinsey & Company(2014). Using customer analytics to boost corporate performance. McKinsey's DataMatics 2013 survey.

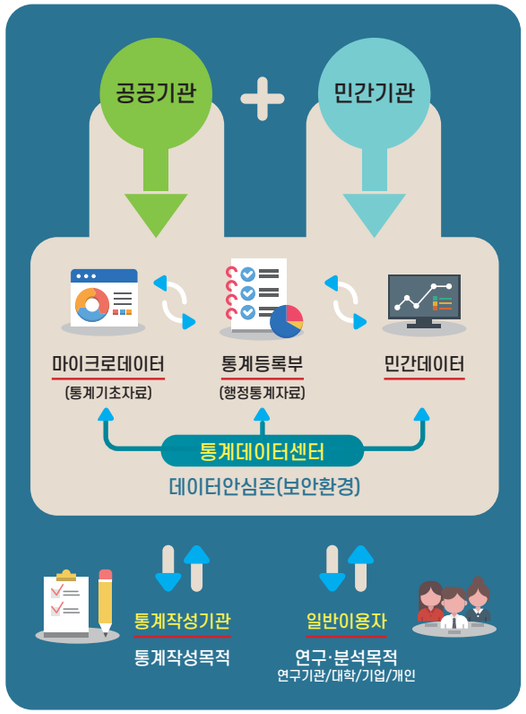
SDC
Statistics
Data Center

행정통계자료와 민간자료를 한곳에!

통계데이터센터 서비스

통계데이터센터가 새로운 서비스로
정보화 사회를 선도합니다.

통계청이 제공하는 승인된 통계기초자료(마이크로데이터),
행정자료를 수집하여 가공한 행정통계자료(통계등록부) 등
통계자료뿐만 아니라 민간자료까지 한 곳에서 연계 · 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

통계청, 정부부처, 지방자치단체, 연구기관 등

모든 기관의 마이크로데이터를 한 곳으로



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

■ 서비스 소개 (2025년 4월 기준)

가. 서비스명 : 마이크로데이터통합서비스(MDIS, mdis.kostat.go.kr)
나. 제공 통계 수 : 172개 작성기관통계 총 382종 제공(통계청 39종)
다. 제공 형태 : 마이크로데이터(조사대상에 따라 인구, 사업체, 가구 기반 자료)

기준		주요 통계
통계청	인구·가구 (16종)	인구동향조사, 인구총조사, 주택총조사, 사망원인통계, 국내인구이동통계, 생활시간조사, 경제활동인구조사, 가계금융복지조사, 가계동향조사, 지역별고용조사, 이민자체류실태및고용조사, 사회조사, 초중고사교육비조사 외 3종
	사업체 (12종)	광업제조업조사, 건설업조사, 전국사업체조사, 기업활동조사, 운수업조사, 서비스조사, 프랜차이즈조사, 경제총조사 외 4종
	농림어업 (11종)	농림어업총조사, 어류양식동향조사, 농가경제조사, 농축산물생산비조사, 어가경제조사, 농작물생산조사, 농림어업조사, 양곡소비량조사, 농업면적조사, 어업생산동향조사 외 1종
통계작성기관 (343종)		전국다문화가족실태조사, 가족실태조사, 노인실태조사, 자동차주행거리통계, 직종별사업체노동력조사, 청소년종합실태조사, 청소년도박실태조사, 국민여가활동조사, 외래관광객조사, 가공식품소비자태도조사 외 333종

■ 서비스 내용

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

나. 서비스 방법
- 다운로드 서비스 : MDIS 포털에서 직접 무료로 다운로드
- 온라인분석 서비스 : MDIS 포털 내 분석 시스템에서 무료로 추출 · 편집 · 분석
- 원격접근 서비스 : 승인 후 이용자가 집 · 사무실 등에서 통계청 서버 접속 후 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
(우) 06097
- 주소 : 서울특별시 강남구 선릉로 612, 6층
- E-mail : MDIS@stat.or.kr

통계청에서 국가통계를 활용하세요!

통계청은 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.kostat.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.kostat.go.kr

지도 위 통계정보 살펴보기

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

통계청
통계인재개발원