

Window of Statistics

통계의 창

2019
WINTER
Vol.24

이슈

데이터 산업, 미래에 다가올 새로운 경제 물결

데이터 경제 활성화, 데이터 거래 시장을 주목하라
나의 데이터에 대한 권리는 누구에게 있는가...마이데이터 산업
지금 바로 통계산업을 육성·지원 해야한다

통계광장

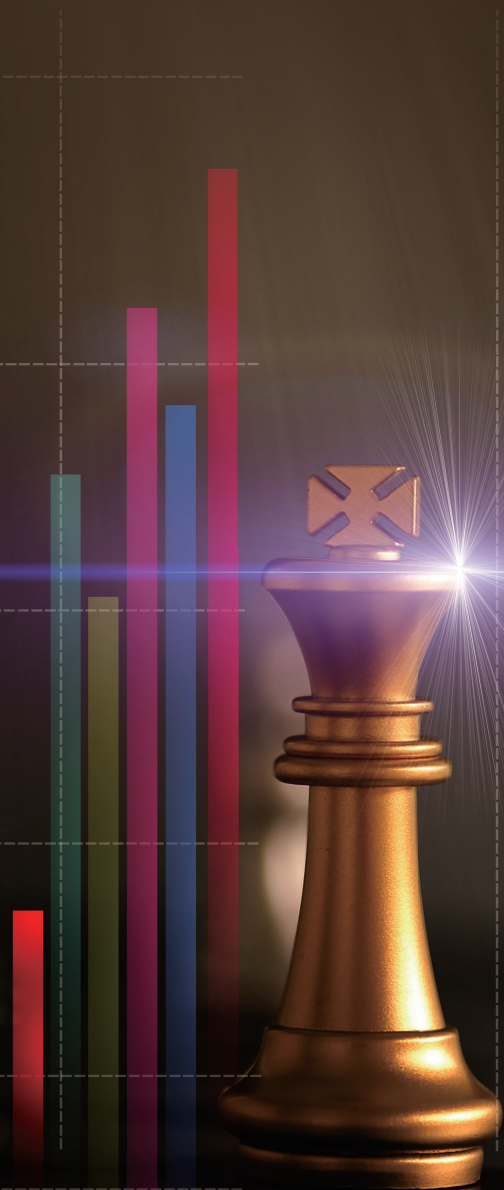
'탈진실'과 가짜뉴스 시대, 통계의 사회적 역할
빅데이터 플랫폼 시대... 데이터 연계체계를 확보하라
한류 콘텐츠의 인기 비결을 알고 싶다면 통계를 찾아보자

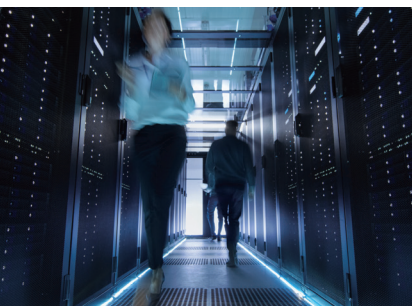
PEOPLE

데이터 경제의 대전환점이 왔다
한국경제산업연구원 김광석 경제연구실장

통계탐방

빅데이터 시대... 가장 핫한 대회가 생겼다!
통계빅데이터 분석 활용대회





ISSUE

002

데이터 경제 활성화, 데이터 거래 시장을 주목하라

— 이재진 / 한국데이터산업진흥원 유통기반실장

010

나의 데이터에 대한 권리는 누구에게 있는가...마이데이터 산업

— 이영환 / 고려대 빅데이터융합사업단 교수

016

지금 바로 통계산업을 육성·지원 해야한다

— 김지영 / 수원대 경영학 교수

통계광장

022

‘탈진실’과 가짜뉴스 시대, 통계의 사회적 역할

— 이광석 / 서울과학기술대학교 IT정책대학원 교수

030

통계에 대한 비판, 그리고 비판에 대한 비판

— 조재근 / 경상대학교 수확응용통계학부 교수

036

데이터 리터러시, 비즈니스 문제 해결을 위한 첫걸음

— 구자룡 / 밸류바인 대표

042

빅데이터 플랫폼 시대... 데이터 연계체계를 확보하라

— 김진철 / 한국정보화진흥원 수석연구원

050

한류 콘텐츠의 인기 비결을 알고 싶다면 통계를 찾아보자

— 김준래 / 통계의창 객원기자

058

지역 자료가 궁금하다면...「e-지방지표」클릭!

— 정현숙 / 통계청 지역통계총괄과 사무관

068

교육

R에 도전하자... 따라가다보면, 나도 R유저®

— 심송용 / 한림대학교 데이터과학스쿨 교수



080

인포그래픽 나도 할 수 있다 series 11

실전에서 자주 사용하는 엑셀 기반의 서열데이터& 발표용 통계 수치 시각화 방법

- 이수동 / 한국인포그래픽협회 대표

084

통그라미, 클릭 한 번이면 나도 통계 전문가! ②

- 정승호 / 영남중학교 교사



PEOPLE

092

데이터 경제의 대전환점이 왔다

- 김광석 / 한국경제산업연구원 경제연구실장

용어 탐구

098

공장에서 나와 현실로 뛰어든 로봇

- 전승민 / 과학칼럼니스트

통계탐방

106

빅데이터 시대... 가장 핫한 대회가 생겼다!

통계빅데이터 분석 활용대회

- 안향숙 / 통계청 통계데이터기획과 사무관



창가의 여유

112

영상 편집... 이것만 알아도 나도 유튜버

- 정영국 / 디자이너

118

당신이 잠든 사이, 몸은 만들어진다

- 김여환 / 가정의학과 전문의(푸른청신경과 원장)

122

직장인도 작가가 될 수 있나요?

- 김용세 / 동화작가

126

통계로 바라보는 세상이야기... 2020 Census 인구주택총조사 실시

- 신동헌 / 통계의창 객원기자

발행일 | 2019년 11월 29일
 발행인 | 임병권
 발행처 | 통계교육원
 기획 | 김정란, 민병희

주소 | 대전광역시 서구 한밭대로 713(월평동) 통계센터 통계교육원
 전화 | 042-366-6161
 팩스 | 042-366-6498
 이메일 | bhmin@korea.kr

디자인 | 동아에스앤씨(02-332-6700)
 인쇄 | 일진인쇄 (02-499-4500)
 ISSN 2005-1379
 ©2019. 통계교육원

※ 『통계의창』에 실린 내용은 필자 개인의 의견이므로 필자의 소속기관이나 본지의 공식적인 견해를 대변하는 것은 아닙니다.





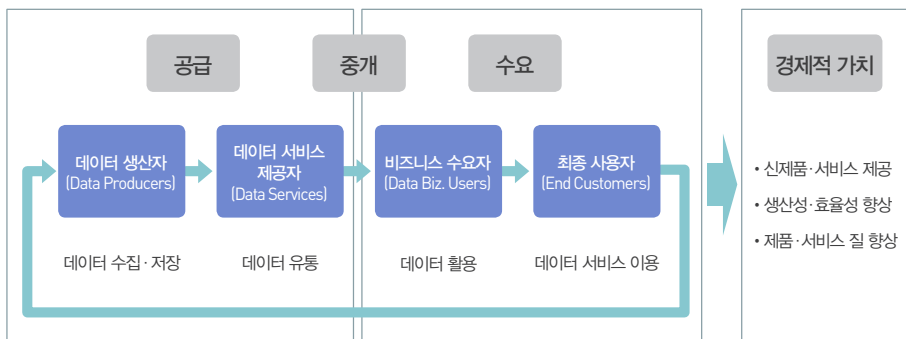
1 데이터 경제 활성화, 데이터 거래 시장을 주목하라

데이터 경제(Data Economy)가 부상하고 있다

4차산업혁명의 시대에 데이터는 흔히 21세기 원유로 비유된다. 데이터가 사람, 자본 등 기존의 생산요소를 능가하는 핵심 자원으로 평가받기 때문이다. 특히 데이터의 활용은 다른 산업을 발전시키는 촉매 역할을 하고, 새로운 제품과 서비스를 창출하는 '데이터 경제'로 이어져 기존 산업의 혁신과 신산업 창출로 대변되는 혁신성장의 중요한 축으로 부상하고 있다.

데이터 경제는 데이터 산업과 이를 둘러싼 데이터 생태계와 직접적으로 관련된다. 데이터 생태계의 가치사슬은 데이터를 생산·수집, 저장·처리, 유통, 활용하는 일련의 과정이다. 이러한 가치사슬을 기반으로 공급-중개-수요 시장을 통해 신제품·서비스, 일자리 등 경제적 가치가 창출되며¹⁾ 이와 관련된 일련의 모든 과정과 연관된 활동은 데이터 산업에 포함된다.²⁾

데이터 경제의 가치창출 체계



출처 : Enter the Data Economy('17, EC), The Rise of the Data Economy('16, IBM) 재구성, 데이터 산업 활성화 전략('18.6, 관계부처합동) 재인용

이러한 데이터 생태계 내에서 데이터가 자원으로 효율적으로 활용되기 위해서는 효율적인 자원 배분이 함께 이루어져야 한다. 현대 경제학에서 가장 효율적인 자원 배분 기제(메커니즘)는 시장이다.³⁾ 따라서 데이터를 하나의 자원으로 보면, 이를 배분하는 데이터 거래 시장을 주목해볼 필요가 있다.

국내 데이터 거래 시장은 확대일로

2019년 한국데이터산업진흥원에서 발표한 '2018 데이터산업 현황조사'에 따르면 2018년 국내 데이터 거래 시장 규모는 3,253억 원 규모이다. 이때 데이터 거래는 데이터 중개·판매와 데이터 신디케이션⁴⁾ 사업을 대표로 영위하는 기업의 매출 규모이다. 다만, 특징적으로 국내 데이터 기업의 데이터 거래 비즈니스의 매출은 이와 같은 데이터 거래 자체 외에도 데이터 가공·분석을 통한 컨설팅 서비스가 상당 비중을 차지하고 있다. 따라서 실제 데이터 거래 시장은 데이터 거래 및 데이터 분석·제공 시장을 포괄한 7,576억 원 규모로 추정해 볼 수 있다.

¹⁾ 관계부처합동, "데이터 산업 활성화 전략", 2018.6.

²⁾ 데이터 산업은 데이터의 생산, 수집, 처리, 분석, 유통, 활용 등을 통해 가치를 창출하는 상품과 서비스를 생산·제공하는 산업으로 정의한다. (2018 데이터산업 현황조사, 한국데이터산업진흥원, 2019.3)

³⁾ 민대홍, 오정숙, "ICT기반 신산업 발전을 위한 데이터 거래 활성화 방안", KISDI, 2018.10.



최근 민간·공공부문 모두 데이터 활용 수요가 증가하고 데이터 기반의 비즈니스가 활발해지면서 데이터 거래 및 분석·제공 시장은 전년 대비 각각 11.5%, 17.1%의 높은 성장세를 나타내고 있어 국내 데이터 거래 시장 전망은 매우 밝은 편이다.

이러한 국내 데이터 거래 시장의 데이터 유통·거래 유형은 크게 플랫폼을 통한 거래와 개별 거래로 살펴볼 수 있다.

데이터서비스 시장 규모

(단위 : 억원)

구분	2014년		2015년		2016년		2017년		2018년(E)		증감률 '17~'18	CAGR '15~'18
	규모	비중	규모	비중	규모	비중	규모	비중	규모	비중		
데이터 거래	2,019	3.5%	2,379	3.7%	2,594	3.9%	2,918	4.3%	3,253	4.5%	11.5%	11.0%
데이터 분석·제공	51,985	90.7%	58,171	90.7%	59,854	90.7%	61,570	90.3%	64,474	89.5%	4.7%	3.5%
정보 제공	3,325	5.8%	3,601	5.6%	3,529	5.3%	3,690	5.4%	4,323	6.0%	17.1%	6.3%
전체	57,329	100%	64,151	100%	65,977	100%	68,179	100%	72,050	100%	5.7%	3.9%

출처 : 2018 데이터산업 현황 조사

4) 데이터 신디케이션은 다양한 데이터를 바탕으로 고객이 원하는 형태로 또는 특정 주제에 따라 구성하여 제공하는 맞춤형 데이터 서비스로 정의한다. (2018 데이터산업현황조사, 한국데이터산업진흥원, 2019.3)

플랫폼을 통한 거래는 운영 주체에 따른 유료 거래와 무료 개방으로 구분할 수 있다. 현재 민간에서 운영 중인 데이터 플랫폼은 자사의 데이터와 외부 데이터를 상품화하여 제공하는데 특정 분야에 국한되고, 공급자 주도로 소극적으로 운영되고 있다. 대표적으로 SK텔레콤의 OPEN API포털/빅데이터허브, LGCNS의 오디피아, KTH의 API스토어, KT의 BIGSIGHT 등이 있다.

공공에서 운영 중인 플랫폼은 정부·공공기관의 개방 데이터를 제공하는 공공데이터포털과 공공·민간데이터의 중개·거래를 지원하는 데이터스토어가 대표적이다. 데이터스토어는 2013년에 최초 구축하여 서비스를 시작하였으며 2019년 10월 현재, 데이터 파일, 응용프로그램 인터페이스(API) 형식의 1,467종의 데이터상품을 제공하고 있다.

이밖에 과기정통부 지원하에 10개 분야 데이터 생산·수집·가공·유통 등을 지원하는 빅데이터 플랫폼이 2019년부터 구축 중이며, 금융, 제조 등 소관 부처에서 지원하는 분야별 데이터 플랫폼도 계획되고 있다. 통계청도 통계자료와 민간자료를 이용하여 연계(가운뎃점)분석하는 데이터 플랫폼인 통계빅데이터센터를 운영하고 있다.

해외 데이터 거래 시장 현황

미국의 주요 9개 기업의 매출은 4억 2,670만 달러

미국은 공개정보, 상업정보, 개인정보 등 다양한 데이터 원천으로부터 데이터를 수집하여 제 3자에게 공유 혹은 재판매하거나 데이터를 가공·분석하여 부가가치 서비스를 제공하는 데이터 브로커 기업⁵⁾을 통해 데이터 거래가 이루어진다. 데이터 브로커 기업은 데이터 전문기업으로 데이터 가공·분석 등을 통해 고객이 원하는 형태의 맞춤형 서비스를 제공한다. 2012년 연방거래위원회(FTC)에서 조사한 주요 9개 업체 매출액만으로도 4억 2,670만 달러(한화 약 4,900억 원)에 달한다.

대표적인 기업으로는 고객데이터 분석 서비스를 제공하는 액시엄(Acxiom) (1969년 설립), 재무정보와 부동산 정보 분석 서비스를 제공하는 코어로직(Corelogic) (1991년 설립) 등이 있다.

중국은 정부주도형 데이터 거래소 운영

중국의 데이터 거래 시장은 2014년 설립된 구이양빅데이터거래소를 시작으로 중국 전역에 개설된 데이터 거래소를 통해 주도되고 있다. 중국의 데이터 거래소는 주로 지방정부의 정책에 따라 설립된 정부주도형 거래소와 데



⁵⁾ "Data Brokers: A Call for Transparency and Accountability", Federal Trade Commission, 2014.



이더 자원을 보유하는 기업이 자발적으로 설립한 데이터거래소로 구성되어 있다. 정부주도형 데이터거래소는 주로 정부의 독자적인 국영기업, 국영기업과 민영기업의 합작형태로 운영되고 있다.

대표적으로 구이양 빅데이터거래소와 상하이 데이터거래소를 꼽을 수 있다. 구이양 빅데이터거래소는 공정한 거래와 시장 신뢰성 확보를 위해 진입 요건(심사-관리-퇴출 등), 거래 보증, 분쟁 조정 등에 관여하는 국가주도형 거래소로 공공데이터를 기반으로 다양한 애플리케이션을 개발할 수 있도록 지원하는데 초점을 맞추고 있다. 이에 보유 데이터 중 공공데이터가 약 80% 수준으로 금융·유통 등 30개 분야 4,200여 종의 데이터를 제공 중이다.

반면, 상하이 데이터거래소는 데이터의 수요-공급을 투명하게 연결하는 중개 역할을 중시하여 데이터 유통·거래 전용 자체 플랫폼(chinadep)을 활용, 온라인 매칭 및 청산·결제 위주로 데이터를 중개한다. 중앙정부, 상하이시정부, 국유기업 등 국유자본(59%)과 민간자본(41%)이 공동출자하여 2016년 출범하였다.



일본의 데이터 거래 시장은 주로 민간 벤처기업 중심으로 유통·거래

일본에서의 데이터 거래 시장은 주로 민간에서 벤처기업들이 중심이 되어, 사물인터넷(IoT) 분야 데이터 유통·거래 플랫폼을 제공하거나 관련 컨설팅 등을 지원하는 형태로 진행되고 있다. 대표적인 데이터 유통·거래 벤처기업으로는 에브리센스재팬(EverySenseJapan)과 제이텍스(J-DEX)를 들 수 있다.

에브리센스재팬은 2014년 설립된 벤처기업으로 IoT 데이터 중개 플랫폼을 구축·운영하고 있다. 플랫폼은 순수 중개(매칭) 기능만 제공하며, 거래 가격은 판매자-구매자가 스스로 결정하는 구조이다. 플랫폼에는 약 3,000개사가 참여하고 있다.

J-DEX는 2016년 설립된 벤처기업으로 데이터 기술정보, 데이터 표준, 연계 정보 등 데이터 거래 지원 컨설팅을 제공한다.

아울러 일본 정부는 데이터 거래 활성화를 위해 2018년 6월 경제산업성 주도로 민간기업들이 데이터 이용·거래 또는 AI 기반 SW 개발·이용 계약시 참고해야 할 가이드라인 3판을 발표하였다. 이는 데이터 거래형태·거래방식 등을 철저하게 사례 중심으로 접근하고 수시로 방법론 등을 진수하기 위함이며 이를 통해 유형별·산업별 특성에 맞는 거래 모델을 도출하려고 노력하고 있다.



데이터 거래가 성립되기 위한 3가지 전제조건

데이터 거래는 공급자(판매자)와 수요자(구매자) 사이에 온·오프라인 방식으로 데이터를 전송·사용·이전하는 행위로 이루어진다. 이러한 데이터 거래가 성립되기 위해서는 다음과 같은 전제조건이 충족되어야 한다.⁶⁾

첫째, 데이터의 가치와 구매력이다. 데이터는 어떻게 활용하느냐에 따라 가치가 상이하며 객관적인 가격산정 기준이 없으나 사고 싶을 만큼 매력적이어야 한다. 아울러 특정 장소에서 특정 형태의 데이터를 공식적으로 이용할 수 있어야 한다.

둘째, 데이터 가공을 위한 품질과 표준이다. 데이터 포맷이 보편적이고 가공이 용이해야 하며, 데이터의 유효성과 활용성 등 데이터 관리 품질이 보장되어야 하며 일정 표준을 준수해야 한다.

셋째, 데이터 소유권 및 개인정보 등 법률문제 해소이다. 다양한 경로로 취득한 데이터를 분석·가공하여 얻은 2차 데이터는 소유권 문제가 명확히 정립되지 않은 상황이다. 이용자 동의가 있더라도 2차 유통시 프라이버시 문제가 다시 발생 가능하며 이를 위해 비식별화 등의 방안이 필요하다.

데이터 거래 시장 활성화를 위해 거래 기반 조성이 필수적

데이터 거래가 확산되기 위해서는 위와 같은 전제조건이 선결되어야 하는데 여전히 쓸만한 양질의 데이터 부족(44%), 불합리한 데이터 가격 산정(37.4%), 데이터 유통 채널 부족(37.4%), 데이터 소재파악 및 검색의 어려움(30.8%), 데이터 품질 문제(23.1%) 등이 데이터 거래 시 애로사항으로 지적되고 있는 실정이다.⁷⁾

따라서 데이터 수집·활용 촉진을 위해서는 데이터 거래 시장 정착이 필수적이다. 이를 위해 해외 주요국은 각자 다양한 방식으로 데이터 거래 활성화를 위한 데이터 거래·유통 체계를 마련하여 운영 중이다. 국내에서도 안전한 거래 촉진을 위해 데이터 거래 기반을 마련하기 위한 다방면의 연구가 진행 중이다.

현재 구상중인 데이터 거래 기반은 다양한 산업의 데이터를 수집·가공한 공급자와 수요자가 만나 데이터를 거래하며 동시에 안전한 거래를 위한 데이터의 품질·표준 관리, 가격산정, 법률검토, 암호화 기술 지원 등의 거래 지원 기능을 수행하는 통합 거래 기반 체계이다.

이를 통해 공급자는 지속적으로 데이터를 수집·발굴하고, 수요자는 더 나은 조건으로 양질의 데이터 상품을 이용할 수 있게 되어 궁극적으로 데이터 거래 시장 활성화로 이어질 것으로 기대한다.

⁶⁾ “데이터 거래 시장(Data Marketplace)”, 한국데이터베이스진흥원, DB Issue Report 제41호, 2014.10.

⁷⁾ 2018 데이터산업 현황조사, 한국데이터산업진흥원, 2019.3

2 나의 데이터에 대한 권리는 누구에게 있는가 ...마이데이터 산업



우리의 모든 생활이 기록, 수집, 저장 되고 있다

4차산업혁명 시대에서 데이터는 선택이 아닌 필수가 되었고 데이터는 21세기 '석유'에 비유되기도 한다. 한국데이터산업진흥원은 데이터 산업의 규모를 약 15조 원(2018년)으로 추정했으



며, 컨설팅 기관 인터내셔널데이터코퍼레이션(IDC)의 연구결과에 의하면 디지털 데이터에서 개인 데이터가 차지하는 비중이 약 70%라고 한다.

우리는 다양한 스마트 디바이스, 전자 장비, 인터넷/모바일 환경에서 소위 빅데이터¹⁾라고 불리는 데이터가 각종 센서를 통해서 수집되고 있다. 스마트폰을 통해서 이루어지는 일련의 개인 생활 활동 과정에서 다양한 데이터가 수집, 저장, 활용되고 있으며 통화, 병원 진료, 쇼핑 등 우리 실생활 전반에 걸쳐 있다.

한 개인이 집에서 병원으로 가는 과정을 살펴보자. 집에서 출발하여 병원에 도착하여 진료를 받는 동안 다양한 행위들이 일어난다. 병원까지 이동하는 동안 스마트폰은 통신 기지국과 끊임없이 신호를 주고받고, 대중교통을 카드로 결제하고, 스마트폰으로 기사 검색과 SNS 서비스를 이용한다. 병원 진료는 대부분 전자 기기를 통해서 이루어지고 진료 내역은 전자문서 형태로 기록되어 저장(Electronic Medical Record, EMR)된다. 우리는 자신도 모르는 사이에 모든 개인 데이터가 생성, 수집, 저장, 공유되는 환경 속에 살고 있다.

이렇게 수집된 데이터의 주체는 누구일까

내가 스마트폰으로 기사검색을 하거나 지인의 소식을 보는 순간 최근에 내가 알아보았던 관광지의 호텔이 추천되고, 더 저렴한 여행 가방과 항공권을 추천받게 된다. 나의 기호와 상황에 딱 들어맞는 상품을 추천받았다는 신기함을 넘어 무섭다는 생각이 든다. 이것은 빅데이터로 수집된 정보를 알고리즘을 통해서 분석 및 추천되는 개인 맞춤형 서비스이다. 그렇다면 나에게 대한 데이터가 이렇게 수집되고 활용되어도 괜찮은 것일까. 나와 관련된 이 데이터의 소유 주체는 누구인가.

내가 통화하고, 검색하고, 구매하고, 진료받아 생성된 개인 데이터는 나한테는 없고 내가 이용한 주체, 즉 통신사, 포털기업, 금융기관, 병원 등 해당 기관의 서버에 고스란히 저장되어 있다. 그렇다면 우리는 개인 데이터를 언제든지 열람할 수 있고, 다운받을 수 있고, 다른 기관에 전달하거나 심지어 우리가 원할 때 데이터를 삭제하거나 폐기할 수 있는 것인가? 그러나 현실은 그렇지 않다. 개인은 데이터를 제공했지만 그 이후부터 배제된 채 기관(기업)들은 개인 데이터를 자신들의 서버에 저장해놓고 자신들의 업무나 사업 목적에만 배타적으로 활용하고 있다.

개인의 동의를 얻어 수집된 개인 데이터의 소유 주체는 누구한테 있는가? 이런 질문으로부터 '마이데이터'에 대한 논의가 시작되었다.



¹⁾ 빅데이터는 단순히 방대하다 거대하다고 정의하지 않고, 빅데이터의 특징을 고려하여 정의된다. 3V를 특징으로 하는데 크기(Volume), 다양성(Variety), 속도(Velocity)이다. 크기는 데이터의 물리적인 크기를 말하며 10테라 이상. 다양성은 정형과 비정형, 속도는 초단위의 실시간 개념을 포함하고 있으나 5G 시대가 되면서 이러한 기준 또한 변하고 있다.

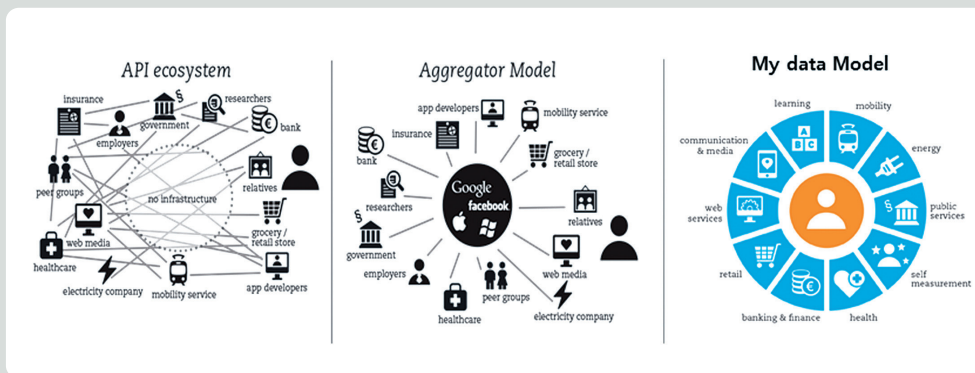
개인의 데이터에 대한 권한의 주체가 개인에게 주어져야

기존에는 스마트폰 통화를 하고 병원 진료를 받고 쇼핑 결제를 하면서 발생된 데이터는 수집한 주체(기업, 기관 등)에 소유권이 있다고 생각해 왔다. 하지만 이제는 개인의 데이터에 대한 소유권이 수집 주체에만 있지 않고 개인에게도 소유 권한이 있다는 새로운 접근 방식이 나타났다.²⁾ 즉, 마이데이터는 데이터에 액세스하고 제어할 수 있는 권한의 주체가 기관(기업)에서 개인으로 이동하는 패러다임이다.

마이데이터에서는 개인 데이터에 대한 접근이 조직이나 기관 중심의 시스템에서 개인과 인간 중심의 시스템으로 변환되고, 개인 데이터는 기술적으로 통제되어야 하며 표준화된 응용 프로그래밍 인터페이스(Application Programming Interfaces, API)를 통해서 개방되고 보호되어야 한다.

〈그림 1〉를 보면 API 에코시스템(API ecosystem)에서는 기업과 기관은 개인 정보를 수집하여 자신의 서비스에 활용했다. 고객 관계 관리(CRM)의 중요성이 부각되면서 통신사, 쇼핑몰, 병원, 공공기관 등 대부분의 기업과 기관에서 개인 데이터와 정보를 수집했다. 이때는 수집 기업과 기관의 경영관리 등 내부 자료로만 활용되었다. 애그리게이터 모델(Aggregator Model)에서는 국내 네이버나 미국의 구글, 애플 등 플랫폼 기업이 등장하면서 데이터 기반의 사업모델이 점차 대세를 이루게 된다. 이들 기업은 개인 데이터를 기반으로 하는 다양한 서비스와 사업 모델을 제공한다. 이전과 다른 점은 경영관리 등의 내부 관점에서 벗어나 데이터를 기반으로 개인에게 직접 서비스를 제공하게 되었다.

그림 1 | 개인 데이터 생태계 변화



출처 : <https://github.com/okfii/mydata>

²⁾ <https://github.com/okfii/mydata#principles>



마이데이터 산업이란 개인이 자신의 데이터에 대한 통제권을 갖는 것

마이데이터 모델(My data Model)에서는 각 개인의 특성을 고려한 '개인 맞춤형 서비스'를 제공한다. 이전에는 플랫폼을 중심으로 데이터의 흐름이 있었다면, 마이데이터 모델에서는 개인을 중심으로 데이터가 이동하고 개인은 데이터에 대한 액세스와 통제 권한이 부여되어 개인이 직접 데이터를 수집, 공유, 활용하는 의사결정을 한다.³⁾ 마이데이터 산업은 의료, 금융, 통신, 유통, 에너지 등 각 산업분야에 있는 개인이 직접 자신의 데이터에 대한 통제권(수집, 저장, 이동, 공유, 활용, 삭제 및 철회 등)을 갖도록 하며 데이터 기반의 개인 맞춤형 서비스를 제공하는 모든 서비스로 포괄할 수 있다.



해외의 마이데이터 산업 추진 현황

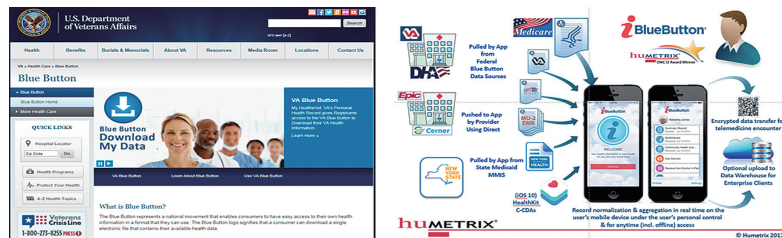
EU는 '일반개인정보보호법(General Data Protection Regulation, GDPR)'을 통해서 개인정보의 보호와 활용에 대한 규정을 마련하여 2018년 5월부터 시행하고 있으며, 영국은 에너지와 통신 분야에서 우선 'Mydata'를 도입하였고 금융분야에서는 'PSD2' 도입과 오픈뱅킹을 통해서 전환점이 되었다. 프랑스는 'Mes INFOS', 'Rainbow Button' 등 파일럿 프로젝트를 중심으로 국민이 마이데이터를 경험하도록 하였으며, 핀란드는 실증 프로젝트를 통해서 마이데이터 모델과 기술적 인프라 등 수립을 시도하고 있다. 미국은 정부의 주도하에 의료, 통신, 교육, 에너지 등의 분야에서 '스마트 공시(Smart Disclosure)'를 통해서 개인 데이터 정책을 추진하고 있다.

구체적인 사례로 미국은 2011년부터 환자가 개인 의료정보를 병원으로부터 직접 다운받거나 다른 병원에 진료 기록을 전달할 수 있도록 소비자의 접근을 가능하게 했다. 미국의 블루버튼(Blue Button) 서비스는 미국 퇴역군인들이 자신이 거주하던 곳에서 먼 곳으로 이사했을 경우 진단서를 발급받는데 물리적인 한계로 인해 불편함을 겪어온 온라인으로 진료 내역을 다운받아서 다른 병원에 전달할 수 있도록 하는 것에서 처음으로 시작되었다. 특히 미국 기업인 휴메트릭스가 개발한 아이블루버튼(iBluebutton)은 환자 개인이 자신의 건강 기록에 대해 접근할 수 있고, 진단 및 복용 약물 목록, 검사 내용, 의사 방문 일정, 입원 내역 등을 열람할 수 있고, 각종 약품, 예방 접종, 알레르기 및 병역을 요약하여 정리해서 보여주기도 한다.



³⁾ 우리나라 금융위원회는 '마이데이터 산업'을 본인 정보의 일괄, 수집 조회 서비스를 기초로 금융 상품에 대한 자문 자산관리 서비스 등으로 정의하고, 정보주체의 개인신용정보 이동권을 보장하고, 그 권리 행사에 기반해 금융회사에게 고객 정보 제공을 의무화시키고자 하는 것이다. 따라서 금융위원회의 마이데이터 산업은 '본인신용정보관리업'으로 그 대상이 분명하다.

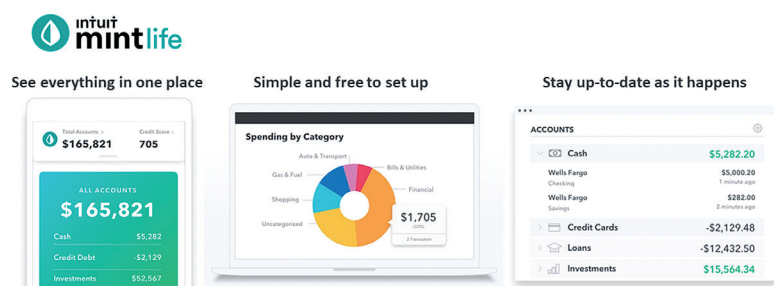
그림 2 | 미국 블루버튼과 휴메트릭스의 아이블루버튼(iBlueButton) 서비스 사례



출처 : VA 홈페이지 <https://www.va.gov/bluebutton>

또 다른 사례를 보면, 미국 기업인 민트(Mint.com)는 개인 맞춤형 금융서비스로서 개인의 자산을 관리해주는 서비스를 제공한다. 이전에 대부분의 금융기관은 자사 고객의 정보만을 이용한 제한된 서비스를 해왔다. 하지만 민트는 개인이 가입되거나 거래하는 금융기관의 정보(ID와 PW)를 받아 해당 기관의 개인 데이터를 대신하여 수집한다. 민트는 개인의 은행 계좌 정보, 증권 계좌 정보, 대출 계좌 정보, 기타 자산 정보 등을 입력하면 통합해서 그래프 등 시각화하여 보여준다. 지출 내역과 소비 패턴 분석을 통해서 자산이 얼마나 증가·감소했는지를 사용자가 일일이 가계부를 기록할 필요가 없이 자동으로 수집되어 제공해준다. 은행뿐만 아니라 증권사, 카드사 등 모든 금융기관의 데이터를 수집하여 한번에 나의 금융 거래 내역을 보여주고, 이를 통해서 맞춤형 금융서비스를 제공하거나 금융 상품을 추천해 줄 수 있다. 내가 현재 사용 중인 신용 카드 및 증권 계좌를 비교해서 추천해주고 이를 통해서 얼마의 자산을 증식하거나 예산을 절약할 수 있는지 알 수 있도록 해주는 서비스이다.

그림 3 | 미국 민트의 금융 서비스 사례



출처 : <https://www.mint.com>



우리나라 마이데이터 산업 발전을 위한 조건

우리나라의 마이데이터 산업 발전과 인간 중심의 데이터 기반의 사회를 위해서는 '마이데이터 포럼'에서 제시한 '디지털 사회 구현'을 위한 접근을 고려할 필요가 있다.

첫째, 사회적 인식변화이다. 개인이 정보 주체로서 데이터에 대한 개인의 권리를 보호하고 개인이 데이터를 통제하고 제어할 수 있도록 해야 한다. 개인은 데이터에 대한 주인의식을 갖고 적극적인 태도로 권리를 행사해야 한다. 이를 위해서는 개인 데이터를 보유한 주체는 개인에게 데이터 소유권을 환원해주는 분위기가 사회 전반에 확산되어야 한다.



둘째, 정부의 제도적 기반 마련이다. 정부가 주도하여 개인의 데이터에 대한 소비자 권한을 강화하고, 개인 데이터를 보유한 기관이 적극적으로 개인에게 데이터와 그 권한을 주도록 제도를 마련해야 한다. 우리나라는 다른 나라에 비해 개인정보보호법 등 강력한 개인정보 정책을 취하고 있다. 개인 데이터 통제권 및 이동권에 대한 원칙과 가이드를 제공하여 중복적인 부분을 통합하거나 조정하는 등의 노력이 필요하다.

셋째, 기업들의 노력과 투자이다. 기업은 개인과 데이터에 대한 신뢰 관계를 우선 구축하도록 해야 한다. 개인 데이터를 독점하지 않고 개인에게 개방하여 개인 데이터를 활용한 다양한 혁신 서비스를 개발하여 개인의 편의성을 높이고 데이터 제공 비용을 낮추도록 해야 한다. 개인 데이터의 보호와 활용을 위한 기술 개발과 인프라 구축에 투자해서 데이터 산업의 수익을 창출하도록 해야 한다. 특히 보안 사고에 대해서는 핵심 소재를 명확히 하여 스스로 책임지는 자세를 보여야 한다.

끝으로 마이데이터 산업은 개인 데이터를 생성하거나 활용하는 모든 분야에서 이중 데이터의 융합을 통해서 제공하는 개인 맞춤형 서비스로 정의할 수 있다. 빅데이터를 기반으로 하는 인공지능 기술의 개인 맞춤형 서비스는 모든 산업 분야에 적용될 것이고 개인과 기관 사이에는 데이터의 권한을 둘러싼 갈등과 분쟁이 증가하게 된다. 따라서 우리는 사업적 기회 창출과 리스크 관리 차원에서 개인에게 정보 주체로서 권한을 부여하는 마이데이터에 대한 이해와 대응을 개인, 기업, 정부 등 모든 사회 구성원들이 함께 노력해야 할 것이다.

3 지금 바로 통계산업을 육성·지원 해야한다

세계적 빅데이터 관련 사업체인 웨스탯(Westat), 닐슨(Nielson), 인테이지(Intage), 매크로 밀(Macromill)의 1인당 생산성은 평균 6만5천 달러로서 삼성전자보다 높으며, 현대자동차의 3만6천 달러의 약 2배 가까운 높은 생산성을 나타내고 있다. 미국기업 닐슨의 경우 11만6천 달러로 삼성전자의 5만7천 달러보다 2배 높은 생산성을 나타내고 있다.

이렇게 빅데이터 시장은 초고속 성장이 예상되고 있으며, 세계는 물론 한국에서도 그 성장세가 신장되고 있는 추세이다. 이제 빅데이터는 새로운 하나의 산업으로 성장할 수 있을 만큼 그 영역을 넓혀 가고 있다. 선진국들은 빅데이터 산업이 활성화될 수 있는 제도적 기반을 구축하고 빅데이터 진흥에 앞장서고 있다. 한국도 빅데이터 산업의 기반인 통계산업의 진흥을 위

한 제도적 장치가 필요한 시점이다. 특히 우리나라의 데이터 사용자들은 기술 수용력이 높고, 우수한 IT인프라를 갖고 있기 때문에 데이터 생산 및 활용이 매우 활발하다. 더욱이 스마트폰 보급이 급격히 확대되면서 이러한 추세는 가속화되고 있다.



통계산업은 무엇일까

2011년 이후, 통계청 및 통계 관련 학회, 업계를 중심으로 통계산업 진흥사업을 추진할 필요성에 관한 논의가 진행되고 있다. 먼저 통계산업이 무엇인지 알아보면 통계산업은 통계데이터를 생산하고, 생산된 통계데이터를 의미 있는 정보로 만들며, 다른 데이터와 연계하거나 가공 또는 분석하여 새로운 부가가치를 창출하는 산업과 이러한 활동을 지원하는 산업을 포함한다고 정의할 수 있다. 통계산업은 통계생산업, 통계정보서비스업, 통계산업지원업으로 분류 가능하다. 또한 통계산업은 범위에 따라 상기 협의의 통계산업과 부분적으로 통계활동을 수행하는 사업체를 포함하는 광의의 통계산업으로 구분할 수 있다.

이러한 통계산업은 고부가가치 산업이며, 관련 일자리는 고임금 일자리이다. 또한 통계산업은 대부분 비교적 최소한의 자원을 사용하고, 공해가 없는 생산물을 만들어내므로 자원절약형 무공해산업이다. 이러한 장점이 있으므로 정부와 업계는 통계산업을 발전시키기 위한 방안을 모색하고 있다.

주요 선진국은「공공·민간 부문간 상승효과 제고를 위한 지침」등 통계분야 활성화를 위한 법령과 제도를 도입하여 적극적으로 통계산업을 육성하고 있는 바, 우리나라도 통계산업 진흥사업을 총괄하는 법적 근거를 마련하고 관련 제도를 도입하는 등 보다 능동적인 대처가 필요하다.



통계산업은 왜 필요한가

통계정보 생산은 인터뷰 및 설문 그리고 자료조사 등을 기반으로 이루어지기 때문에 자원의 이용 및 공해를 유발할 가능성은 매우 낮다.

통계생산을 위해 드는 비용은 인건비를 비롯한 운영비와 패널 구성비용 및 조사비용 등이 대부분이다. 자동차 및 가전의 경우 인건비보다 자원 투입비용이 큰 부문을 차지하지만, 통계생산을 위해서는 인건비와 자원투입 비용이 큰 차이를 보이지 않거나 인건비가 더 높은 경우가 다수이다. 따라서 통계산업은 고부가가치를 발생시킬 수 있는 고도화



사회의 필수적 산업부문이라고 할 수 있다.

이러한 측면에서 4차 산업혁명, 빅데이터 시대를 맞아, 각종 데이터를 활용하여 새로운 부가가치를 창출하는 통계산업 육성은 국가발전의 신(新)모멘텀 발굴 및 좋은 일자리 창출 측면에서 반드시 필요하다.

통계산업을 기반으로 한 빅데이터 산업은 정보통신기술(ICT)산업의 발전과 더불어 상호보완 관계를 가지고 성장하고 있다. ICT산업은 다양한 분야에서 통계정보를 바탕으로 한 산업영역을 넓히고 있으며 이 과정에서 통계 관련 산업을 확장시키는데 기여하고 있다. ICT산업은 앞으로도 지속 성장이 가능한 산업으로 빅데이터를 중심으로 한 통계산업의 파급효과를 더욱 높일 것으로 예상된다.



주요국의 통계산업 현황

미국의 통계산업 규모는 2015년중 명목GDP의 0.10% 이다. 그중 정부부문과 민간부문 비중이 각각 14.2%, 85.7% 이다. 미국 정부는 중앙부처 및 지자체가 민간 통계회사에 통계사업 발주를 큰 규모로 하고 있다. 주요 통계조사업체로는 엠피알(Mathematica Policy Research, MPR), 웨스텍, 닐슨 등이 있다.

일본 통계산업 규모는 2015년 기준 명목GDP의 0.04% 수준이다. 일본 통계산업에서 민간부문은 85% 내외, 정부부문은 15% 내외 이다. 해외리서치 업체들은 일본에 상당수 진출하여 있다. 2011년 세계 매출 5위안에 있는 업체인 닐슨, 칸타(Kantar), IMS 헬스(IMS Health) 지에프케이(GfK) 등이 모두 일본에 자회사나 지사를 보유하고 있다. 일본 업체들의 해외 매출은 2015년 현재 약 39개 기업체가 47억 엔으로 전체 매출의 2.7%에 불과하다. 주요 통계업체로는 인테이지, 매크로밀 등이 있다.

영국 통계산업 규모는 2015년 기준 명목GDP의 0.16% 정도이며, 2009년 이후 연평균 5.3%가 증가한 것으로 나타났다. 민간부문과 공공부문의 통계산업 비중은 각각 94.2%, 5.8%이다.

영국 조사업계는 세계화를 선도하고 있는 것으로 평가되는데, 영국은 세계 70여 개 국가의 조사기업들이 참여하는 마켓리서치소사이어티(Market Research Society) 등의 세계적인 조사기업협회를 보유하고 있다.

독일 통계산업 규모는 2015년 기준 명목GDP의 0.086% 이다. 독일 통계산업 중 공공과 민간부문 비중은 각각 20% 내외, 80% 내외이다. 주요 통계업체로는 지에프케이가 있다.





국내 통계산업의 주요 현황

국내 통계산업 규모는 광의 기준으로 2015년중 1조 6,758억 원으로 GDP의 0.11%로 추정된다. 또한 통계산업은 2011년부터 2015년까지 연평균 5.9% 증가하여, 같은 기간 중 GDP 평균증가율(4.3%)보다 높은 증가율을 보였다. 우리나라 통계산업에서 공공부문 비중은 약 40% 내지 50%, 민간부문은 50% 내지 60%이다. 선진국의 경우 통계산업에서 민간부문이 차지하는 비중이 대체로 70%를 상회하고 있다.

국내 통계산업의 주요 업종별 현황을 살펴보면 다음과 같다.

조사업종인 국내 시장조사 및 여론조사업의 2015년중 매출액은 7,934억 원, 종사자수는 8,184명으로 추정되었다. 2014년 기준 국내 매출은 85.0%, 해외 매출은 15.0%로 나타났다. 국내 매출액 중 의뢰기업이 순수 한국기업인 경우는 71%, 다국적기업과 해외에서 의뢰받은



조사는 각각 11.2%와 2.8%이었다. 국내매출 중 공공기관과 기업매출 비중은 각각 20%, 80% 정도로 추정된다. 2014년 기준 마케팅조사와 여론조사 매출비중은 각각 62.0%와 38.0%로 나타났다.

국내 대다수 중소 조사통계기업은 기술 개발 및 인력 확보의 어려움을 겪고 있다. 소규모 시장에서 저가 입찰 및 가격 경쟁력 중심의 영업활동을 수행하고 있어, 부가가치 생산성 및 기업 성장성이 낮은 것으로 파악되고 있다.

한국조사협회에 가입된 41개 기업 중 상당수의 대형기업이 해외기업의 법인이나 지사이다. 세계에서 가장 큰 조사업체인 닐슨의 현지법인인 닐슨코리아, 일본 매크로밀 한국법인인 엠브레인(Embrain), 칸타TNS의 현지법인인 칸타코리아 등이 있다. 국내기업으로는 한국리서치, 코리아리서치, 한국사회여론연구소 등이 있는데, 이중 한국리서치를 제외하고는 대부분 매출 100억 원 이하의 소규모 기업이 대부분이다.

통계산업을 영위한다고 볼 수 있는 금융업종 중 신용정보사와 신용카드사는 많은 정보를 보유하고 있다. 금융부문에서 개인의 정보는 개인정보보호법과 신용정보의 이용 및 보호에

관한 법률에 따라 보호를 받기 때문에 현재 금융자료를 타 산업에 활용하는 것은 모두 불법이 된다.

통신업종의 경우 지방자치단체의 관광효과 분석을 지원하거나, 금융과 통신의 결합형태로 생활금융 플랫폼 서비스를 제공할 수도 있다. 이동전화는 위치 추적이 가능하므로 개인정보 보호 문제와 밀접하다. 방송통신위원회 등은 「개인정보 비식별 조치 가이드라인」을 2016년에 발표하였다. 그러나 업계에서는 동 가이드라인은 데이터 사용으로 문제가 야기되면 결국 데이터를 사용한 기업이 책임을 지게 되어 있어 개인정보를 충분히 활용하기에는 한계가 있다고 보고 있다. 시민단체도 개인정보보호 측면에서 「개인정보 비식별 조치 가이드라인」을 반대하고 있다.

포털업종은 데이터 가공을 통한 비즈니스 모델을 별도로 운영하고 있지 않다. 회원으로 가입된 이용자들의 다양한 정보는 수입을 목적으로 가공하거나 별도로 데이터를 판매하고 있지 않으며, 특정 시점의 주제들을 가공하여 데이터 포털(데이터랩)에 무료로 제공하거나, 특정 주제에 대한 분석결과를 무료로 제공하고 있다. 한편 포털의 데이터 판매업무로는 주요 검색어 관련 업무를 일부 하고 있으나, 동 사업은 전체 영업의 매우 작은 부분에 불과하다.

의료보건업종의 경우 건강보험심사평가원이 운영하고 있는 보건 의료 빅데이터는 심평원 내에 빅데이터센터를 설치하여 건강보험 데이터베이스(DB)를 활용한 다양한 통계분석이 가능하도록 지원하고 있다. 빅데이터센터에는 이용자가 직접 데이터 센터에 방문하여 건강보험DB를 이용하거나 원격분석시스템을 이용하여 건강보험 자료를 분석할 수 있도록 하고 있다.



국내 통계산업진흥법 제정 필요

우리나라의 취약한 통계산업의 발전을 위해서는 여타 육성 필요 산업의 사례와 같이 통계산업진흥법 제정을 통한 지원이 시급하다. 이와 함께 공공정보의 상업적 활용 가능성 강화, 개인정보보호 체계 개선, 국가통계 마이크로데이터 서비스 개선, 통계인력 양성 및 통계정보 관리 강화, 통계조사업체의 대형화 및 전문화 등의 조치가 시급하다. 특히 개인정보보호로 인한 장애요인 극복 방안 마련이 중요한바, 개인정보를 보호하는 가운데 개인정보가 포함된 데이터의 이용을 활성화하기 위해서 「개인정보보호법」상 개인정보 이용을 필요한 보안조치를 강화하되 허용범위를 명확히 하고, 동 허용범위를 확대해나가는 조치를 적극적으로 검토할 필요가 있다.

이러한 통계 산업의 발전은 올바른 국가정책 수립 및 건전한 여론형성을 위한 통계 품질향상에도 크게 기여할 것이다.





‘탈진실’과 가짜뉴스 시대, 통계의 사회적 역할

미디어가 내보내는 뉴스는 보통 사실들 가운데 가장 객관적인 지표가 되어야 하고, 사실과 진실에 다가가기 위한 중요한 판단 기준이어야 한다. 하지만 오늘날 이도 옛말이 됐다. 미디어들은 특종과 속보 과열 경쟁으로 사실 확인 없이 상호 베끼거나 정보의 출처마저 철저히 다루지 못하면서 잘못된 사실들을 전달하거나 오류로 가득 찬 뉴스를 생산하는 일이 비일비재 하다.

무엇보다 인터넷 환경이 열리고 어느 누구나 제 목소리를 낼 수 있는 뉴스 생산 문화가 되면서 영향력 있는 누리꾼들이 퍼 나르는 소문, 뒷담화, 가짜정보 등이 이들 대중매체와 뒤섞이고 뒤엉켜 더 큰 오류가 만들어지는 거짓 세계를 생성하고 있다. 게다가 특정의 알고리즘 학습 기제를 통해 인공지능이 뉴스를 자동 생산할 수 있게 되면서 ‘딥페이크(deepfake)’라 불리는 가짜 이미지와 영상 가공물이 범람하고 자동 생성되는 가짜뉴스 ‘봇(bot)’들이 활개 칠 공산이 더 커졌다.



가짜 뉴스가 범람하는 이유

각종 이미지, 정보, 영상 데이터에 의한 대중 왜곡이 범람하고 그러다 보니 이에 대응해 ‘팩트 체크’가 일상인 현실을 우린 살아가고 있다. 사회적 사건에 대한 노이즈(소음)의 왜곡 데이터들이 증가하면서 이를 설명하는 뉴스와 정보들은 주어진 사건과 역사에 대한 사실과 진실 값을 얻기 위한 콘텐츠로 기능하기 보다는 사실과 진실을 흐리는 사악한 ‘터미(허수)’로 기능한다. 이들 데이터의 자의적 왜곡이 문제인 것은 우리의 의식을 호도해 사실상 이전에는 크게 도전받지 않던 진실 값을 계속해 뒤흔들고 있기 때문이다. 가짜는 진실이 존재하고 진실이 밝혀지더라도 진실의 공신력에 위해를 입힌다. 그것이 가짜 기호로 무장한 오정보와 가짜뉴스들이 범람하는 이유이기도 하리라.

대중은 이들 과잉 정보와 가짜뉴스로 인해 피로감을 호소하게 되고, 진실에 대한 판단 자체를 대부분 유보하게 된다. 현실로부터 진실을 찾는 행위를 쉽게 포기하는 것이다. 오늘 ‘탈진실(post-truth)’의 목표는 바로 여기에 있다. 즉 탈진실을 옹호하는 이들은 우리 대부분이 현실로부터 기꺼이 행하는 진실 찾기 행위에 스스로 무기력하게 되고 포기 상태로 이르는 것을 진정 바란다.



탈진실의 효과

옥스퍼드 영어사전의 정의에 따르면, ‘탈진실’은 “여론을 형성할 때 객관적인 사실보다 개인적인 신념과 감정에 호소하는 것이 더 큰 영향력을 발휘하는 현상”을 뜻한다.¹⁾ 사실의 진위와 상관없이 진실을 무력하게 만들고 외려 특정 신념이나 감정이 영향력을 발휘하도록 한다는 것은, 무엇보다 대중 의식을 끌어들이л 흡인력이 없으면 불가능하다. 사실상 그 흡인력이란 것은 대중매체의 온갖 강렬한 자극의 ‘황색 저널리즘’에서 올 수도 있지만, 대개는 우리 주위 뉴스원의 권위에 대한 신뢰에서 유래한다.

국가 통치자의 트윗 메시지, 주류 대중 매체 언론 보도, 메이저 포털 뉴스, 소셜미디어의 인플루언서, 스타 유튜버 등이 가장 큰 대중 흡인력을 지닌 출처이자 권위의 근원이다. 아이러니하게도 이들은 권위의 출처이지만 탈진실의 탄생지이기도 하다.

먼저 탈진실의 출처는 변방의 누군가 이름 없는 이들의 거짓 소문들로부터 시작된다 기보다는 이렇듯 말과 정보의 독점적 권위를 갖고 이를 확산할 메가폰을 쥔 이들에 의해 주도된다. 권위적 통치자의 입, 가짜 댓글 알바부대, 무분별한 언론 속보 과열 경쟁, 포털의 연예뉴스와 댓글, 글로벌 기업들의 후원 뉴스 등 사회 엘리트의 언설이 포스트진실의 중요한 근원이다. 이들은 기후위기를 부정하기 위해, 전쟁의 명분을 만들기 위해, 각종 사회 비리를 감추기 위해, 역사의 진실을 은폐하기 위해 가짜와 페이크 유사과학이나 사실들을 대거 유포한다.²⁾ 그들의 목소리가 효과적인 이유는 이를 마치 ‘바이러스 전염(바이럴)’마냥 실어 나르는 대중의 빠른 전달력이다. 오늘날 ‘소셜’미디어의 가치는 인간관계의 재구성 가운데서도 특정 정보나 사실을 빠르게 전달해 거대한 정서적 교류를 이어 나가는 능력에 있다. 탈진실 효과는 이렇게 소셜 테크놀로지를 통해 쉽게 번식하고 엘리트 권력의 여론 통제에 중요한 요소로 활용된다.

둘째, 가짜뉴스는 우리에게 ‘초현실(hyperreality)’ 효과를 더한다. 즉 “리얼리티에 더 가까이 다가간 것 같을수록 영상들은 그만큼 더 흐릿해지고 더 흔들린다.”³⁾ 사건들을 애



1) 리 매킨타이어, 『포스트 트루스: 가짜 뉴스와 탈진실의 시대』, 김재경 역, 2019, 두리반, 19쪽 재인용.

2) 부시 행정부의 이라크전의 명분부터 트럼프 행정부의 기후협약 탈퇴에 이르기까지 통치자의 말이 어떻게 진실과 무관한 말들이었는지에 대한 비판적 평가는, 미치코 가쿠타니, 『진실 따위는 중요하지 않다: 거짓과 혐오는 어떻게 일상이 되었나』, 김영선 역, 2019, 들베게 참고.

3) 히토 슈타이얼, 『진실의 색』, 안규철 역, 2019, 위크루, 14쪽 인용.

기하는 수많은 영상과 이미지 정보나 뉴스들은 마치 사태의 정확성을 얘기하는 듯 보이지만, 오히려 우리가 '진실의 색'을 읽을 수 있거나 판단할 수 있는 능력을 갉아먹는 효과를 낸다. 빅데이터 시대에 뉴스와 정보는 폭발 지경이지만, 반대로 인간의 판단 능력이 점점 더 약화되고 상황 파악이나 인식 능력이 점차 고갈되는 형국이다.

셋째, 이와 같은 초현실을 더욱 부채질하는 데는 기술적으로 '필터버블(filter bubble)' 효과가 그 배경에 있다. 필터버블은 맞춤형 데이터에 익숙해져 그것의 과잉 정보 수취가 이루어지면서 각자가 편향된 정보 거품에 갇히게 되는 효과이다.⁴⁾ 소셜미디어에서 특히 이런 현상이 강화된다. 내 주위의 '소셜' 인적 네트워크에 연결된 이들이 비슷한 성향을 유지하면서 스스로 무오류성의 착시 감옥에 갇히고, 다른 판단 가능성을 배제하는 일이 흔해진다. 알고리즘의 현실은 이 필터버블 효과를 더 극대화한다. 개인이 지닌 성향에 따라 자동화된 맞춤형 예측 서비스를 제공하는 인공지능 알고리즘 기제는, 현대 인간이 소비하는 대부분의 데이터와 정보 콘텐츠를 개인 취향의 소소한 세계로 가둔다. 진실이 흐려짐과 동시에 거기에 도달하기 위한 정보의 도관(導管)도 얇아지고 납작해지는 형국이다.

마지막으로, '실재의 소멸' 효과가 궁극적으로 발생한다. 탈진실, 초현실, 필터버블의 3중 효과는 모든 역사적, 인본적, 사회적 가치들의 자명한 질서를 불완전하고 비결정적인 지위로 만들어버린다. 우리가 알고 지내던 분명한 듯 보이는 실재가 저 멀리 달아나는 것이다. 그러다 보니 오늘날 상징 권력은 특정의 가치와 담론을 자명한 질서로 내세워 억지로 강요하기보다는, 혼돈 속 여럿 가짜들을 기술적 알고리즘으로 자동 생성하거나 댓글알바를 고용해 만든 가짜 터미 속에 진실의 가치를 뒤섞는데 골몰한다. 이와 같은 거짓과 허구는 우리의 진실처럼 군림하지만 실제 어떠한 소통도 행하지 못하면서 계속해 우리가 시도하려는 현실의 비판적 인식을 방해한다.

탈진실의 가짜뉴스 시대에는 어찌 보면 오정보와 노이즈(소음)를 대거 발산하는 쪽에 승산이 있다. 이를테면 누군가에 대한 흑색선전이 법리적으로 '근거 없다'는 법적 판단에도 불구하고 계속해서 비상식적으로 비방과 악플이 계속해 진행되는 것은 이와 같은 연유에서다. 결국 데이터 과잉과 오정보의 질서는 특정의 사안에 대한 진실이 저 멀리 사라지고 수많은 다른 가짜 해석들에 대중들을 노출시키면서, 어떤 사안에 대해 우리 스스로 사색하는 것을 불안정하고 어렵게 만드는 데 그 목적을 지닌다.



4) 엘리 프레이저, 『생각 조종자들: 당신의 의사결정을 설계하는 위험한 집단』, 이현숙·이정대 역, 2011, 알키.



가짜뉴스는 여론과 정서의 조정과 변조 방식에 큰 관심

데이터과잉과 가짜뉴스가 대중의 진실에 대한 판단을 포기하게끔 하면서 결국 엘리트 권력은 우리 마음속 깊숙이 가라앉은 욕망, 정서, 선호, 취향 흐름의 조절과 통제에 직접 개입하고자 한다. 정보 과잉의 질서 속에서 이성적 판단이 어려워지면, 그다음 대중이 판단할 수 있는 선택지는 본능과 직관일 수밖에 없다. 합리적이고 비판적 추론에 따른 진실 찾기 과정이 점차 힘들어지면서, 진실의 지위는 흐릿해지고 대중은 진실이나 사실보다는 직관과 감성적 판단에 친숙해진다.

인터넷에서 수없이 흐르는 감정과 정서의 흐름과 패턴들에 관심을 지닌 오늘날 엘리트 권력은, 대중 여론 이전의 여론, 즉 입소문과 비공식적 '뒷담화'를 집요하게 관리하는 수준에 이르렀다. 동시대 권력은 대중 정서의 흐름을 이리저리하게 너울거리게 하고 여론과 정서의 물꼬를 바꾸는 조정과 '변조'(modulation) 방식에 큰 관심을 보인다. 가짜와 탈진실 정보의 범람과 이의 특정한 변조는 대중 정서적 차원에서 또 다르게 엘리트 권력이 미치는 범위를 넓히는 새로운 기법이 되어가고 있는 것이다. 오늘날 인공지능과 결합된 알고리즘 기계 장치는 순수한 데



DEEP FAKES CHEAP FAKES

이더공학의 오류 없는 숫자의 세계인양 포장하면서도, 동시에 수많은 가짜와 거짓 정보들을 퍼뜨리며 판독 불가능한 현실 질서를 공고히 한다.

최근 인공지능 기술은 거짓과 탈진실을 방조하거나 이를 강화하는 동영상 기법에 응용되면서, 가짜 이미지와 동영상 제조에 줄곧 쓰이고 있다. 가령, ‘딥 페이크’(deep fakes)와 ‘싸구려 페이크’(cheap fakes)가 무한 확대되는 현실을 보라. 싸구려 페이크는 말 그대로 포토샵 작업 등으로 실체가 아니라는 느낌이 남아있는 아마추어 이미지나 영상이라면, 딥 페이크는 인공지능의 딥러닝 기술을 활용해 실제 사람 얼굴의 일부를 변형하거나 합성해 만들어 가짜와 실제의 경계가 거의 사라진 경우다. 딥 페이크는 ‘진실보다 더 진실한’ 거짓 혹은 가짜의 고해상도 이미지나 영상인 경우가 흔하다. ‘팩트 체크’를 통한 가짜들의 공식 판별 방식은, 아직까진 인공지능 봇들에 의한 가짜 이미지와 영상의 자동 생성에 비하면 진실게임 싸움에서 양적으로 밀리고 질적 판단 능력에서 소박하기 그지없다.



또 다른 가짜뉴스의 문제는 사회적 소수에 집중

‘증거의 정치학(the politics of evidence)’적 견지에서 봐도, 정교한 페이크 상징들은 점점 더 사회에서 일반적으로 수용되던 사실 ‘증거’ 능력을 소멸시키는 대신 불확실한 현실을 주조하는 측면이 크다.⁵⁾ 자칭 진짜라 주장하는 딥 페이크들로 인해 현실은 점점 불투명해지고 생성되는 이들 가짜정보 양의 방대함으로 인해 참고할 진실의 정확성은 점점 희박해간다. ‘탈진실’의 또 다른 문제는 이 정교한 거짓 조작과 혐오 대상들이 주로 여성, 인종소수자, 난민 등 사회적 소수에 집중된다는 데 있다. 가짜에 휘둘리는 표적들이면서 가짜들이 노리는 주된 제물들이 바로 현실 속 힘없는 이들이란 점을 기억해야 한다.

아이러니하게도 이 모든 가짜와 탈진실의 질서를 작동시키는 수면 아래에는 첨단 기술네트워크와 인공지능 알고리즘을 기반으로 해 현실의 모든 것을 예측하려는 인공지능과 데이터과학의 최첨단 과학의 세계가 있다. 대중 의식은 스펙터클과 혼돈의 탈진실로 향하는 반면, 정보 권력의 최첨단 세계는 빅데이터 분석을 통해 과학적 투명성을 추구한다. 가짜뉴스에 의해 크게 영향받는 대중의 불투명한 현실 세계와 달리, 반대로 엘리트 통치 영역은 그들을 비추는 투명한 세계에 기초한다.

언제 어디서든 개별 시민들의 신원을 투명하게 확인하고픈 권력 욕망은 오늘날 안면인식 기술과 생체 기술이나 개인 신체 프로파일링의 극한 수집 도구를 창안해낸다. 지문, 홍채, 생체 안면 인식, 신체 동작, 신체 이식 칩의 알고리즘 자동화는 불



⁵⁾ Paris, Britt & Joan Donovan, *Deepfakes and Cheap Fakes: The Manipulation of Audio and Visual Evidence*, Data & Society, New York, 2019.

투명한 현실로부터 정보 권력의 자리를 굳건히 보장받고 사물의 질서를 보다 예측 가능하게 하는 일이기도 하다. 탈진실 시대야말로 정교한 신체 데이터 프로파일링 관리를 통해 통치 권력을 유지하려는 욕구가 더 크게 작동한다. 우리가 대체로 사건과 진실의 올바른 이면을 보지 못하는 불투명한 상태에 놓인다면, 반대로 엘리트 계급은 우리 각각을 투명하게 비추는 스크린을 원한다. 가령, 이를 위해 중국은 14억에 육박하는 국민 모두를 관리하기 위해 '사회신용체계'(social credit system)를 도입하고 그들의 개인차를 안면 인식으로 확인하고 24시간 관리하는 빅브라더 구상을 진행하려 한다. 결국 가짜와 거짓 뉴스로 인해 대중의 진실 이해력은 퇴화하지만 반대로 권력이 우릴 비추는 능력은 계속해 확장된다. 오늘날 자본주의의 탈진실은 결국 서로 상반된 듯 보이는, 진실의 불투명성 대 개별 신체의 투명성이란 쌍두마차에 의해 유지된다.



가짜뉴스 시대 점점 중요해 지는 통계의 사회적 역할

최근 비정규직 노동 인구 실태 파악과 관련한 통계 자료가 공개됐다. 이로부터 이제까지 수면 위에 잡히지 않던 비정규직 노동자 숫자가 크게 늘었다. 몇 가지 관련 통계 질문들이 추가되면서 통계치에 잡히지 않았던 노동 인구가 더 추가된 것이다. 물론 결과적으로 잘 된 일이다. 왜냐면 부정확한 통계값을 바로잡아줌으로써 실제 비정규직 노동자 숫자를 명확하게 시민에게 알리게 됐기 때문이다. 물론 이는 고용 상황에 민감한 정부의 고용 성과 홍보에 상당히 불리하게 작용할 수도 있다. 하지만, 여러 우려에도 불구하고 우리가 잘못 알고 있는 현실의 교정과 진실에 좀 더 다가갈 수 있도록 통계 오류를 바로잡았다는 점에서 충분히 긍정적인 의의를 지닌다.

전통적으로 통계에서 제시되는 숫자는 특정 편견의 개입과 데이터 왜곡이 일어날 수 있는 영역으로 고려해 다들 의심을 거두지 않았다. 반면, 이번 노동 통계 사례는 가짜 뉴스가 범람하고 진실과 가짜를 구분하기 어려운 시대에서 통계값이 지닐 수 있는 사실에 다가갈 수 있는 힘 또한 증거하는 경위로 볼 수 있다. 즉 갈수록 사회에 점점 널리 퍼지는 편견과 오류를 제거하고 사안의 객관적 정황 증거를 제시하는 출구로서 통계의 사회적 역할을 확인할 수 있었다. 더군다나 이번 통계치로 인해 한 사회 비정규직 노동자들의 보다 정확한 실태와 그 지형을 확실히 인식할 수 있게 됐다는 점에서 향후 통계의 사회적 역할이 보다 선명해졌다고도 볼 수 있다. 예컨대 사회적으로 잘 파악되지 않는 약자들의 그림자 노동과 복잡한 고용 계약으로 인해 그 숫자가 제대로 잡히지 않거나 분류가 잘못된 노동층에 대한 오류를 걷어내는 작업에 통계의 역할을 찾을 수 있는 것이다. 다시 말해, 거짓





과 혐오가 일상이 되더라도 그 반대편에서 사회적 약자와 타자가 놓인 비가시적 영역들을 지속적으로 찾아 드러내는 보편 통계 과학적 접근이 강조된다.

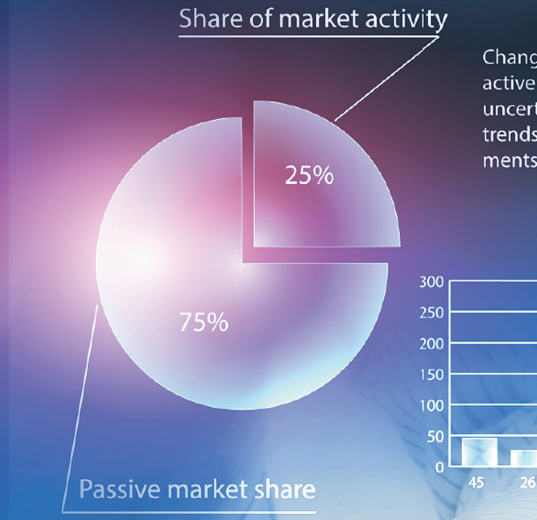


궁극적으로 시민의 데이터 리터러시 능력의 향상 필요

수없이 증가하는 가짜들 앞에서 이를 판별해내는 팩트 체크 전문 언론기관이나 통계 과학의 검증 역할은 점점 중요해진다. 하지만 하루에도 수없이 발생하는 가짜 데이터의 공급을 막기에 이들 기관의 대응은 현실적으로 수적 한계에 봉착할 수밖에 없다. 궁극적으로 시민의 ‘데이터 리터러시(문식력)’의 향상 없이는 향후 대중의 현실 판단력이 흐려질 확률이 높다. 시민들 자신이 오늘날 복잡한 가짜의 바다들로 인해 파생되는 심층의 관계들을 파악할 수 있는 최소한의 지식이나 통제력 혹은 비판적 판단 능력을 배양하도록 해야 한다. 예를 들어 기존의 규격화된 문화와 데이터 소비를 벗어나서 데이터와 숫자에 대한 비판적 인식과 시민 자신의 지혜를 확장할 수 있는 데이터 리터러시 교육 논의가 사회적으로 본격화되어야 한다. 가짜와 탈진실에 대적할 ‘데이터 시민권’적 사회 설계 논의 또한 필요하다. 그럴 때만이 온갖 가짜뉴스와 오정보의 늪에서 허우적거리는 우리 자신을 구할 수 있다.



Projected sales of main products in 2013

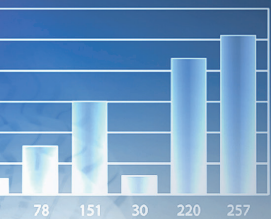


통계에 대한 비판, 그리고 비판에 대한 비판

알바생은 취업자일까 실업자일까

사회, 경제통계 강의실에서 “지금 우리나라의 실업률은 얼마쯤 될까요?”라는 질문을 던져보곤 하는데 학생들로부터 흔히 듣게 되는 답은 “30%, 40%, 50%” 등이다. 그 질문에 이어 저녁이나 주말마다 알바 일을 하고 있는 학생을 찾아 자신이 취업자라고 생각하는지 물어보기도 한다. 그러면 학생들은 무슨 말도 안 되는 질문을 하느냐는 표정을 짓거나 심지어 불쾌한 반응을 보이기까지 한다. 몇 년째 커피가게에서 일을 하거나 저녁마다 편의점에서 일하고 있다 하더라도 알바 일은 취업해서 하는 일과 전혀 다르다는 것이다. 학생들이 생각하는 실업률이란 일정한 연령 이상의 국민 가운데 이른바 정규직으로 일하

es in the activity of the
and passive market is
tain. Established positive
in various market seg-



지 않는 사람의 비율을 뜻하는 듯싶다. 그런 정의에 따르자면 알바 일을 하는 사람은 취업자보다는 실업자에 더 가까울 듯하다.

그런데 과연 통계청에서도 알바 학생을 그렇게 볼까? 전혀 아니다. 대학생이면서 알바 일을 하는 사람이라면 본인의 생각과 상관없이 고용통계에서 엄연한 취업자로 잡힌다. 게다가 실업률이라는 것이 오천만 국민 가운데 실업자의 수를 계산한 것도 아니다. 15세 이상인 사람 가운데 육아 또는 가사 활동을 하는 사람 그리고 학교나 학원 같은 데에서 공부만 하는 사람 등은 아예 경제활동을 하지 않는 인구로 간주된다. 즉 실업자도 아니고 취업자도 아닌 것이다.

경제활동인구란 일자리를 가진 사람과 일자리를 찾는 사람을 뜻하므로 취업자와 실업자를 합한 인구에 해당한다. 그런 분류에 따르자면 알바 일을 하는 학생은 고용통계에서 취업자가 된다. 실업자로 인정받으려면 일할 의사와 능력이 있어야 함은 물론 구직활동까지 해야 하니 공식 통계에서 실업자로 인정받기가 쉽지 않은 것이다. 결국 실업률은 실업자 수의 증감에 따라서만 변하는 것이 아니고 다양한 범주로 구성된 비경제활동인구의 변동에 따라서도 값이 달라질 수 있다.

통계의 객관성, 정확성이란 무엇인가

통계청에서는 표본 집단을 대상으로 그런 내용을 조사하기 때문에 실업률 값은 대개 학생들이 생각하는 값의 1/10 수준에 머문다. 그런 통계를 알고 나면 “통계청에서 실업률을 낮추려고 꼼수를 부려서 현실과 동떨어진 통계를 만들고 있는 것 같다”고 말하는 학생도 있다. 물론 그럴 리야 없다. 수입을 얻기 위해 일주일에 한 시간만 일하면 취업자로 보는 것은 우리나라 통계청이 마음대로 정한 것이 아니고 국제노동기구(ILO)의 규정을 따른 것이다. 그런데 국제노동기구는 왜 실업률을 그렇게 정의한 것일까? 그런 지표들이 현실을 제대로 반영하는 것일까? 혹시 다른 정의를 상상해볼 수는 없을까? 일에 대한 생각이나 노동시장의 사정이 나라마다 많이 다를 텐데 국제기준을 꼭 통일시켜야 할까?

학생들은 이렇게 해서 ‘통계의 객관성, 정확성이란 무엇인가?’라는 질문과 대면하게 된다. 측정하고 숫자를 헤아리기 전에 엄밀한 기준과 정의에 따른 세밀한 분류기준이 먼저 필요하다는 것, 그리고 세상을 실업률처럼 단 하나의 통계 속에 가둘 수 없다는 것을 이해하고 나면 이제 통계를 공부할 준비가 된 것이다.

통계는 숫자가 아니고 정보이다

통계는 근대사회의 산물이다. 구미 각국에서 정부기관이나 민간단체들을 중심으로 각종 통계를 만들기 시작한 것은 19세기 전반기의 일이었다. 흔히 오늘날을 ‘빅데이터의 시대’라고 부르기도 하는데 데이터에 대한 열정을 생각해보면 19세기 전반기를 최초의 데이터시대라고 불려도 될 정도다. 무엇보다 당시는 산업혁명(1차 산업혁명)의 시기로서 정치, 산업 그리고 사람들의 생활 등 여러 면에서 급격한 변화가 일어난 때였다. 그 극심한 혼란의 와중에서 대중의 불만을 헤아리고 체제를 유지하기 위해서는 노동자들의 생활



상, 도시빈민 문제, 그리고 범죄나 자살 등의 문제를 객관적으로 파악해서 적절한 대책을 마련하는 것이 시급했다. 인류 역사상 처음으로 다양한 통계데이터가 절실하게 필요한 시대가 온 것이다.

1830년대 영국에서 지역마다 민간이 주도하는 통계단체들이 생긴 것도 영국이 가장 먼저 산업혁명을 겪은 나라였기 때문이다. 그런 단체 중 하나가 1834년에 생긴 런던통계협회[오늘날의 왕립통계협회(Royal Statistical Society)의 전신]다. 그 단체를 만든 사람 가운데에는 『인구론』으로 유명한 맬서스, 프로그램이 가능한 컴퓨터라는 개념을 처음 제시한 배비지와 같은 당대의 명사들이 들어있었는데 그들의 면면을 살피는 것만으로도 통계에 대한 당시의 관심과 열정을 넉넉히 짐작할 수 있다.

처음에 그 협회는 밀 짚단을 묶은 그림을 상징으로 삼았는데 자세히 보면 그 밀 짚단을 묶은 끈에 “Aliis exterendum”이라는 라틴어 문장이 보인다. 영어로는 보통 “To be threshed out by others”라고 옮기고 “우리 단체는 자료를 모으기만 할 뿐 해석은 하지 않는다”라는 뜻으로 읽는다. 19세기 전반 사람들이 볼 때 통계의 생명은 최대한 현실을 있는 그대로 조사해서 데이터로 모으는 데에 있었기 때문이었다. 따라서 당시에는 이론이나 주장의 영향을 받지 않는 순수한 통계만이 객관적인 과학이라는 대접을 받을 수 있었고 데이터에 이론을 개입시키고 이런 저런 해석을 덧붙이는 것은 통계를 오염시키는 금기사항으로 간주되었다.

그런 순수성이 오래 지속되었을까? 당연히 얼마가지 못했다. 19세기 중반이 지나자 런던통계협회는 겨우 20년 만에 그들이 상징으로 삼았던 밀 짚단 그림을 살짝 바꾸게 된다. 밀 짚단을 묶은 끈에 적혀있던 라틴어 문장을 슬그머니 빼버린 것이다. 그 이유는 실제 조사활동을 통해 이론이 없는, 해석이 없는 통계란 불가능하다는 사실을 알게 되었기 때문일 것이다. 예나 지금이나 통계 조사원들이 조사를 나가면 실업자들이 ‘나는 실업자입니다’라고 써 붙이고 있는 게 아니다. ‘있는 그대로’ 헤아려달라고 기다리고 있는 순수하고 깨끗한 데이터는 어디서도 찾기 어려웠다. 누구를 실업자로 분류할 것인지, 무엇을 기준으로 삼아 빈곤을 측정할 것인지, 또 질병이나 사망원인을 어떻게 분류할 것인지 등등 사람들이 만드는 모든 통계에는 자못 치열한 이론적, 논쟁이 필요했고 타협이 필요했을 것이다. 그리고 국제적인 협력이 필요했을 것이다.

그리하여 굉장히 이른 시기인 1850년대부터 각국의 통계공무원들과 전문가들이 격년으로 국제통계회의(International Statistical Congress)를 열어 정보를 교환하고 국제적인 통계 기준을 마련하기 시작했던 것이다.

통계에 대한 신뢰와 불신

19세기가 지나는 동안 통계는 점점 중요해졌고 그에 따라 나라마다 통계 전담 정부 기관을 만들고 인력을 양성하는 등 국가 경영에 필요한 통계를 만드는 데 많은 노력을 기울이게 되었다. 그에 따라 다양하고 복잡한 통계들이 많아지면서 데이터에 대한 해석도 점점 까다로워졌다. 한때 단순해보이고 사실을 그대로 드러낸다고 여겨지던 통계라는 것이



영국 런던통계협회의 두 가지 문장



이제 많은 공부가 필요한 전문가의 일이 된 것이다. 그러다 보니 실업률처럼 일상을 사는 사람들이 생각하는 것과 꽤 달라 보이는 정의를 가진 통계들도 많이 생기게 되었고, 이런 까닭에 국가기관에서 만드는 통계 역시 열광과 신뢰의 대상이 되기도 하고 때로는 불신의 대상이 되기도 했다.

그 결과 오늘날까지도 우리는 대표적인 거짓말의 하나로 통계를 꼽아야 한다는 말을 자주 듣고 있다. 이런 상황이다 보니 각 나라들은 통계조사에 필요한 기법과 이론, 데이터 해석 방법 등을 개발하는 것은 물론 정부기관에서 만든 통계에 대한 신뢰를 얻기 위해서도 각별히 노력을 기울이게 되었다. 아무리 열심히 만들었다 한들 사람들이 믿어주지 않는다면 그 통계의 생명은 끝장이기 때문이다.

우리나라는 어땠을까? 한국에서 근대적인 인구조사가 처음 실시된 것은 1925년부터였다. 서양보다 백 몇 십 년 늦게 센서스가 실시되었는데 그마저도 우리가 준비해서 조사한 것이 아니고 조선총독부가 시작한 것이었다. 이후 식민지 시절, 그리고 해방과 전쟁을 거치는 동안 우리나라가 체계적인 통계를 생산할 인력을 키우고 조직을 갖추기는 어려웠을 테다. 그러다가 1990년대 초에 통계청이 만들어지고 얼마 뒤에 우리나라가 경제협력개발기구(OECD)의 회원국이 되면서 통계도 국제 기준에 맞추어 급성장하게 된다.

통계청이 승격 30주년을 맞는 오늘날 한국 통계는 다른 나라의 전문가들이 배우러 올 만큼 부쩍 성장했고 통계에 대한 신뢰도 크게 높아졌다.

그렇다면 이제 통계청은 누구나 언제든지 믿고 찾아보는 기관으로 성장했을까? 아쉽게도 아직 충분한 정도는 아닌 듯하다.

정부기관들이 만든 홈페이지를 찾아보면 다들 맨 앞에 보도 자료를 내는 창을 마련해 두고 있다. 그런데 유별나게도 통계청 홈페이지에는 보도 자료와 더불어 ‘해명 자료’를 보여주는 곳이 있다. 주로 신문, 방송, 정치권 등에서 통계청을 비판하는 주장을 내놓으면 그에 대해 해명하는 내용이 들어있다. 그곳에 등장하는 비판의 제목과 내용들을 훑어보자. “통계에 물 타는 통계청”, “통계에 대한 불신을 자초하는 통계청”, “정권의 눈치를 보는 통계청”, “엉터리 물가지수를 발표하는 통계청”…… 여전히 ‘통계란 조작하기 쉬운 것이므로 의심해볼 필요가 있다’는 생각이 사라지지 않고 있는 듯하다. 물론 사회나 경제통계는 원래 다양한 해석이 가능한 것이므로 언론이나 정치권의 날카로운 비판과 감시는 우리나라 통계와 통계청을 위해서도 당연히 필요할 것이다.



통계는 정치를 평가하는 하나의 도구이다

그런 비판에 대한 통계청의 ‘해명’은 대개 “통계청은 국제기준과 전문적인 판단에 따라 통계를 만들고 해석하고 있을 뿐 정치와는 무관하다”는 것이다.

혹시 있을지 모르는 외부의 입김을 막기 위해 통계청의 위상을 더 높이고, 전문성을 살릴 수 있도록 책임자의 임기를 보장하는 등의 정치적 독립은 물론 필요할 테다. 그러고 나면 통계가 정치로부터 자유로워질까? 그런데 통계청은 정말 그래야 할까? ‘정치적 독립’이란 과연 무슨 뜻일까?

우리는 5년마다 대통령을 뽑고 4년마다 국회의원 선거를 한다. 기후변화 문제를 심각하게 인식하고 대책을 마련하겠다는 공약을 내세운 후보나 정당이 선거에서 승리해 정권을 잡았다고 해보자. 마땅히 통계청은 기후변화와 관련된 통계를 더 정비하고 이전까지 만들지 않았던 새로운 통계도 개발해야 할 것이다. 사회적 약자들의 편에 서겠다고 약속한 후보가 대통령이 되어 정책을 펴면 통계청은 그 정책의 효과를 잴 수 있는 새롭고 정교한 지표를 만드는 것이 마땅할 테고, 따라서 통계의 기준은 변화를 주어야 하고 새로운 통계도 개발해야 한다. 시대변화를 반영하기 위해서는 물론이고 후보의 공약을 보고 인물을 선택한 국민의 의사를 받들고 권력자들이 공약을 제대로 이행하는지 감시하기 위해서도 통계가 필요하기 때문이다. 이런 의미에서 통계는 당연히 정치와 무관할 수 없다.

그런 까닭에 통계청에 대한 시민과 언론, 정치권의 비판도 업그레이드되어야 하지 않

을까? 잠시 다른 정부기관과 비교해보자. 국가기관들 가운데 통계청처럼 사실을 비틀어 거짓말을 일삼는다고 욕을 먹는 곳이 또 있을까? 일기예보를 잘못했다고 가끔 기상청을 비난하는 경우가 있지만 이런 식은 아니었던 것 같다. 당연히 기상청 홈페이지에는 ‘해명 자료’를 따로 올려두는 곳조차 없다. 통계청을 비판하는 언론 기사나 정치권의 목소리는 대개 “통계청이 어떤 의도를 갖고 통계를 비틀거나 감춘다.”고 주장한다. 그런데 일기예보가 맞지 않으면 기상청이 어떤 이유에선가 ‘예측에 실패’했다고 하지 숨은 의도를 갖고 왜곡해서 예보했다고 비판하는 사람은 아무도 없다. 비판을 하더라도 적어도 기상청에서 일하는 사람들의 전문성은 인정하고 비판을 하는 셈이다. 그런데 통계청에 대한 비판에서는 그렇지 않다. 통계청 사람들은 종종 전문성은 물론 공직자로서의 직업윤리도, 시민의식도, 책임감도 없는 사람들 대접을 받는다.

이제 업그레이드된 비판이 필요할 때이다

물론 새삼 말할 필요도 없이 기상학은 과학이지만 통계는 과학이면서 또한 정치다. 그러므로 통계청의 통계에 대해 왜곡, 은폐, 누락, 눈치 보기 통계라고 비판하는 것도 통계청이 늘 새겨들어야 마땅할 것이다. 그런 비판 덕분에 통계는 맷집만 키운 것이 아니라 점점 과학적인 이론과 방법으로 무장하고 더욱 정교해졌기 때문이다. 하지만 언론 등으로부터 통계가 변함없이 거의 비슷한 비판만 받는 것은 아무래도 아쉽다.

지난 10월에 세계경제포럼(WEF)에서 발표한 국가경쟁력순위에 따르면 한국은 세계 140여 개 나라 가운데 열세 번째였다. 통계가 선진화되지 않고 이처럼 한국이 선진국 대열에 올라서는 것이 가능했을 리 없다. 만약 세계 여러 나라 통계청의 활동을 순위 매긴다면 우리나라 통계청의 순위가 국가순위보다 낮지는 않을 것이란 말이다. 지금쯤이면 이처럼 훌쩍 성장한 우리나라 통계와 통계청에 대한 보다 더 업그레이드된 비판을 기대할 때도 되지 않았을까?

그런데 그런 새로운 비판 기능을 지금까지처럼 계속 언론이나 정치인들에게만 맡겨둬도 될지 모르겠다. 관심과 비판은 다양해질수록 질이 높아지는 법, 비판의 수준을 높이기 위해 통계전문가들, 대학의 통계학과에 있는 교수들이 지금보다 더 적극적으로

나서면 어떨까? 사실 대부분의 통계학 전문가들은 통계청이 하는 일이나 통계청이 겪는 곤혹스런 상황들에 대해 관심도 없고 아는 바도 없다. 통계와 관련이 깊은 사회학이나 경제학, 사회복지학, 정치학, 공중보건학, 행정학 등의 전문가들도 크게 다르지 않을 테다. 그들의 무관심이 지속된다면 통계에 대한 비판 역시 왜곡, 누락, 정치적 편향이라는 오래된 테두리 안을 계속 맴돌지 모른다. 통계나 통계청에 대한 우리의 생각은 여러 전문가들이 다양한 지식과 경험을 동원해서 늘 지켜보고 적극적으로 목소리를 내는 분위기 속에서 함께 쑥쑥 성장할 것 같다. 그럴 때 통계는 지금보다 더 다양한 생각을 가진 사람들이 서로 비판하고 어우러지는 흥미진진한 과학과 정치의 마당이 될 수 있을 것이다.





데이터 리터러시, 비즈니스 문제 해결을 위한 첫걸음



도대체 뭐가 문제야

빅데이터를 가지고 있는데 이것을 어떻게 사용해야 할지 모르겠다는 기업이 있다. 반대로 연일 뉴스에서 빅데이터 관련 이야기를 듣게 되는데 그런 데이터가 없어서 걱정이라는 기업도 있다. 데이터가 있어도 걱정, 없어도 걱정이 앞서는 시대에 데이터를 전문적으로 다루지 않는 임직원들은 가시방석에 앉아 있는 기분이다. 젊은 직원들은 그나마 R이나 파이썬과 같은 데이터분석 도구를 배울 기회와 능력이 어느 정도 있다.

하지만 컴퓨터도, 인터넷도, 스마트폰도, 별도로 사용방법을 배웠던 임원들은 불안할 수밖에 없다. 특히 의사결정을 하는 과정에서 데이터를 적극적으로 활용하라는 지시를



받은 상태다. 기본적으로 데이터가 가지고 있는 의미도, 수집하고 분석하는 방법도, 그리고 그 결과를 활용하여 통찰하는 방법도 익숙하지 않았던 아날로그 세대에게는 디지털로 전환되고 있는 경영 환경에서 데이터와 관련된 스트레스가 커지고 있다.

데이터 리터러시 즉, 데이터를 수집하고 분석하고 해석하는 역량에 한계가 있어서 업무에 애로가 생기는 상황이다. 이런 현상은 데이터에 집착하기 때문이다. 특히 의사결정을 해야 하는 임원들은 익숙하지 않은 디지털 환경에서 데이터 과학자들이 전문적인 용어와 분석 방법으로 빅데이터를 설명하는 장면을 보면서 한계를 느끼는 것이다. 실무자들 역시 이 부분은 임원들과 다름이 없다. 그런데 잘 한번 생각해 보면 이는 앞뒤가 바뀐 것이다. 데이터와 데이터 분석이 필요한 이유는 우리가 가지고 있는 어떤 문제를 해결하기 위한 것이다. 데이터 분석만을 위해 데이터가 필요한 것이 아니다. 데이터와 관련해서 보면 수단과 목적이 뒤바뀐 것이다.

비즈니스 환경은 끊임없이 새로운 의사결정을 요구한다. 일상적인 의사결정은 자동화 되어 있기도 하지만 중요한 의사결정은 아직 대부분 의사결정자의 의사결정에 의해 이루어진다. 이때 데이터의 도움을 받으면 보다 쉽게 통찰할 수 있고 오류를 줄일 수 있다는 생각으로 데이터를 먼저 찾는다. 우리가 데이터를 가지고 있으면, 데이터에서 멋진 통찰을 할 수 있을

것이라고 기대를 하게 된다. 과연 그럴까? 혹은 데이터가 없기 때문에 통찰을 하지 못한다는 불평을 하며 책임을 회피하고 싶어 한다. 이 경우에 데이터를 제공하면 멋진 통찰을 할 수 있을까?

중요한 것은 데이터가 아니라 해결해야 할 문제다. 문제가 무엇인지 정의하는 것이 먼저다. 예를 들어, 시장에 수많은 스낵이 있고 여러 회사에서 치열하게 경쟁을 하고 있다. 서로 점유율 싸움을 하는 완전경쟁시장이다. 그런데 우리 회사에 새로 부임한 사장은 시장을 확대하고 매출을 증대하고 싶어 한다. 여기서 문제는 무엇일까? 경쟁이 문제일까? 점유율이 문제일까? 시장을 확대하고 매출을 높이하고자 한 사장의 방침이 문제일까? 아니면 데이터가 없는 우리 회사의 시스템이 문제일까?

문제의 본질을 정확하게 정의하는 것이 무엇보다 중요하고 먼저 고민해야 진짜 문제를 해결할 수 있다. 여기서 문제는 경쟁이 치열하지만 정체된 스낵시장을 돌파할 새로운 제

품이 없다는 점이다. 데이터는 이 문제를 해결하는데 필요한 많은 요소 중 하나일 뿐이다. 회사 내부에는 매출 데이터도 있고, 과거 진행했던 시장조사 데이터도 있고, 관능조사(맛 테스트)의 결과도 있다. 아직 신제품 관련 조사를 한 적이 없기 때문에 직접적인 데이터는 당연히 없다. 이 시장에서 우리는 선발자도 아니고 시장지배력을 가지고 있는 것도 아니다. 그동안 여러 번 신제품을 만들어 보기도 했지만 실패를 많이 해서 선뜻 나서는 마케터가 없다. 여러분이 의사결정자라면 이러한 상황을 돌파하기 위해 무엇을 할 수 있을까?

다른 예를 하나 들어 보자. 정부나 지방자치단체에서 비무장지대(DMZ)가 현재 어떤 이미지가 있고 앞으로 어떤 브랜드가 되면 좋을지 궁금해 하고 있다. 이 가정은 일반기업에서 브랜드의 이미지나 브랜딩을 위해 궁금해 하는 상황과 동일하다. 문제가 무엇일까? DMZ에 대한 브랜드 연구를 위해서는 우선 문제를 명확하게 해야 한다. DMZ를 관광목적지 브랜드로 규정하고 관광목적지를 찾고자 하는 이해관계자들의 인식 속에 어떤 이미지가 형성되어 있는지 그리고 브랜드 정체성으로 어떤 가치요소를 고객의 인식 속에 구축해야 차별적인 포지셔닝이 되는지를 알아야 한다. 즉 DMZ의 이미지와 브랜드 정체성 요소를 파악하는 것이 연구의 문제가 된다. 이를 바탕으로 강력한 브랜드를 구축하는 데 도움이 되는 자료를 만드는 것이 연구의 목적이 된다.



데이터는 어디에 있지

먼저 스낵시장에 대한 데이터는 어디에 있을까? 정해진 시장을 돌파하기 위한 전략으로 차별화된 신제품을 개발하여 고객에게 새로운 가치를 제안하고 이를 통해 스낵시장에 새로운 바람 일으키기 위해서는 시장의 특성을 정확하게 이해해야 한다. 스낵시장에서 기업들은 어떤 종류와 맛으로 시장을 나누고 있는지 고객은 이러한 구분에 동의하는지 아니면 원하는 제품이 없어서 마지못해 기업이 제안하는 제품을 구매하고 있는지 파악해야 한다. 이를 위해 스낵의 맛지도를 만들면 어떨까?

스낵시장을 맛으로 구분하기 위해서는 제품 범주별로 어떤 맛이 있는지 알아봐야 한다. 우리 회사에서 생산하는 제품만이 아니라 시장에서 판매되고 있는 전체 제품에 대한 분석이 필요하다. 이 문제를 해결하기 위한 데이터는 우리 회사 내부 보다는 실제 시장에 있다. 시장조사가 필요한 이유이다. 슈퍼나 마트나 편의점 등에서 기초적인 데이터를 수집하고 제조사별 판매액을 수집하거나 추정하여 범주별 맛의 비중을 계산하여 정리하면 하나의 표나 그림이 완성된다. 표본조사가 적합하다. 이때 데이터는 현재의 시장을 그대로 보여줄 수 있는 실제 시장에서 수집하고 일부 2차 자료를 찾아서 보완할 수 있다.

그렇다면 DMZ에 대한 이미지를 파악할 수 있는 데이터는 어디에 있을까? 대체로 DMZ를 알고 있거나 방문한 경험이 있는 사람들은 DMZ라고 하면 전쟁, 분단, 북한, 군인, 평화, 생태 등 여러 단어를 제시할 수 있을 것이다. 그런데 이를 구체적으로 확인하려고 하면 이런 데이터를 수집해 놓은 곳을 찾기 어렵다. 일반적인 사람들의 인식 속에 어떤 단어들이 들어 있기는 한데 우리가 직접 조사를 하지 않으면 원시데이터(raw data)나





분석된 결과를 얻을 수 없다. 정책을 추진하는 차원에서 본다면 구체적인 결과물이 아닌 상상의 단어를 근거로 의사결정을 할 수는 없다. 이럴 때 가능한 데이터는 구조화된 설문 조사를 통해 직접적으로 수집하는 1차 자료이다.

그리고 온라인상에 있는 소셜미디어나 뉴스에서 언급되는 단어를 수집하여 텍스트 마이닝에 의한 빅데이터 분석으로 이미지를 파악할 수도 있다. 네이버 뉴스검색의 결과를 RSS 피드로 엑셀의 파워쿼리를 이용해 크롤링 할 수 있다. 또는 빅카인즈에서 뉴스검색을 할 수 있으며 특정의 조건을 설정하여 필요한 데이터 중심으로 크롤링 할 수 있다. 크롤링한 결과는 엑셀 파일 형태로 저장할 수 있어서 다음 분석에 활용할 수 있다. 경우에 따라서는 구글트렌드 같은 빅데이터 분석 사이트에서 어떤 단어로 검색을 했는지에 대한 빈도와 연관어 등을 알 수 있는 서비스를 제공받을 수도 있다.

따라서 데이터는 그냥 있는 것이 아니다. 이것을 아는 것이 중요하다. 문제를 해결할 수 있는 적절한 데이터를 필요한 양만큼 수집해야 한다. 물론 구매 이력 데이터나 고객 행동 데이터와 같이 실시간으로 시스템에 의해 자동으로 수집되는 데이터도 있다. 이 데이터는 데이터 마이닝을 통해 구매확률을 계산하여 다음 프로모션에 활용하여 보다 높은 매출 성과를 올릴 수도 있다. 즉, 어떤 문제가 있다면 그 문제를 해결할 수 있는 데이터는 따로 있으며 필요에 따라 가장 적합한 데이터를 수집해야 한다. 데이터가 있어서 분석하는 것이 아니라 비즈니스의 문제를 해결하기 위해서 필요한 데이터를 수집하고 수집된 데이터를 분석하는 것이다.



왜 통계를 들먹이는 거야

연말이 가까워지면서 기업의 임원들은 지난 1년간 투자한 마케팅 비용에 대비해서 제대로 마케팅 활동을 했는지, 그 성과가 어떤지 궁금해 한다. 마케팅 성과 지표로 매출액과 이익을 기준으로 분석하기도 한다. 광고도 하고 판촉도 하고 이벤트도 하고 가격할인도 한다. 그런데 이런 활동들이 과연 투자 대비 효과(ROI)가 있는지 궁금한 것이다. 그리고 대부분의 기업에서는 핵심성과지표(KPI) 중에서 고객 추천 지수(NPS)와 고객 만족도(CSI)를 매우 중요하게 관리한다. 우리 회사를 추천하는 추천 의향율과 고객의 만족 정도를 매년 1회 이상 설문조사를 통해서 조사하고 그 결과를 바탕으로 NPS와 CSI를 산출한다. 나름 체계적인 마케팅 성과 시스템이 구축되어 있다. 그런데 갑자기 의문이 들기 시작했다. 마케팅 예산을 전년 대비 대폭 증액을 했는데 목표한 매출액과 이익을 내지 못하고 있다. 최고경영자가 마케팅 ROI를 구해서 원인을 찾아오라고 한다. 마케팅 성과가 기대한 만큼 나오지 않은 원인이 무엇일까? 어떻게 하면 원인을 찾아 문제를 해결할 수 있을까?

그럼 마케팅 성과와 통계는 무슨 관계일까? 통계를 몰랐어도 지금까지 데이터를 보는데 문제가 없었다. 과연 그럴까? 단순히 보면 투자 대비 수익인 마케팅 ROI는 쉽게 계산이 된다. 수익에서 투자금을 빼고 이 금액을 투자금으로 나누어 주면 된다. 문제는 수익이다. 어떤 금액을 수익으로 볼 수 있느냐 하는 것에서 논쟁이 많이 일어난다. 마케팅 부서에서 관리하는 금액으로 영업이익이 있다. 영업이익은 매출이익에서 판매와 관리비를 빼준 값이다. 판매와 관리비는 대체로 마케팅 비용이라고 할 수 있다. 일견 타당한 것 같지만 순이익이 아니기 때문에 과대 계산될 여지가 높다. 만약 수익을 특정할 수 있다면 마케팅 ROI를 계산하는 것은 쉬운 일이다. 상대적인 비교가 가능하기 때문에 매우 유용한 지표가 된다. 그러나 실제 마케팅 활동에 대한 수익을 추정하기 어렵기 때문에 아직도 지표로 활용되기에는 제약이 많다.

여기에 매출에 영향을 미친 변수로 무엇이 있는지 살펴봐야 한다. 광고, 판촉, 이벤트, 가격할인, 그리고 신제품 수, NPS, CSI 등 많은 변수들이 매출에 영향을 미칠 수 있다. 이 영향력을 정확하게 추정해야 어떤 마케팅 활동을 할 때 단위당 효과가 극대화 되는지 알 수 있다. 이를 추정할 수 있는 방법으로 민감도 분석이 있다. 통계분석은 회귀분석을 사용한다. 종속변수인 매출액에 영향을 미치는 독립변수에 따라 단순회귀분석 혹은 다중회귀분석을 사용한다. 회귀계수를 구하여 독립변수의 영향력을 바탕으로 어느 변수가 더 많은 영향을 미치는지 파악할 수 있다.

여기에 필요한 데이터는 어디에 있을까? 제품별 혹은 품목별 매출액은 재무팀에 있고, 신제품 수는 마케팅팀에서 관리하고, 광고와 이벤트 비용은 프로모션팀에서, 판촉과 가격할인은 영업팀에서, 그리고 NPS와 CSI는 고



객만족팀에서 주로 관리하고 있다. 사내에 흩어져 있는 데이터를 수집하고 모델링을 해야 하는 마케팅 성과 분석 담당자는 통계에 대한 이해 없이 분석할 수 없다. 이 결과를 받아든 관리자나 임원들 역시 기본적인 통계를 알아야 의미를 해석할 수 있다. 단순히 이야기해서 다중회귀분석을 하려면 독립변수들 간에 상관성이 없어야 하고(다중공선성), 종속변수와 독립변수는 계량자료여야 하며, 민감도 분석을 위해서는 각 변수에 대해 최소한의 표본수를 가지고 있어야 한다. 더 중요한 부분은 독립변수를 무엇으로 할지와 그에 따른 독립변수의 측정값을 모두 가지고 있어야 한다는 점이다.

만약 고객 만족도와 고객 추천 의향을 간에 어떤 관계가 있는지 궁금하다면 이 두 변수 간에는 상관관계가 있는지 확인해야 한다. 상관 분석이 필요하다. 상관 분석의 통계치로 상관계수를 구하게 되며 상관계수가 0.5 이상이면 상관이 있다. 만약 두 변수 간의 상관계수가 0.7 이상이면 강한 정의 상관이라고 한다. 즉, 만족도가 높아지면 추천의향도 높아지는 것으로 판단할 수 있다.

그리고 추천 의향이 연령 집단에 따라 차이가 있는지 알아보기 위해서는 분산분석(ANOVA)을 통해 집단 간 평균 차이를 검증해야 한다. 만약 분산분석을 통해 구한 P값(p-value)이 0.042가 나왔다면 유의수준($p < 0.05$) 값보다 작기 때문에 집단 간에 차이가 없다는 귀무가설을 기각하게 된다. 즉, 유의수준 5%(신뢰수준 95%)에서 연령 집단에 따른 추천의향에 차이가 있다는 결론에 이르게 된다. 최소한 기초적인 통계치를 이해하고 측정 척도와 분석방법을 알고 있어야 제대로 의사결정에 반영할 수 있다.



데이터에 대한 실질 문맹률을 낮추는 첫걸음

디지털 시대에 데이터를 기반으로 비즈니스를 전개하는 것은 너무나 당연한 흐름이다. 처음 적응하는데 어려움이 있지만 조금만 있으면 데이터로 문제를 해결하는 데 익숙해진다. 한국인의 문맹률(illiteracy)은 세계에서 최저 수준이다. 거의 제로에 가깝다. 그러나 OECD(경제협력개발기구)의 문해력 보고서에 따르면 한국인의 '실질 문맹률'은 75%로 OECD 22개 국가 중 최하위 수준이다. 문자는 읽지만 문장의 의미를 이해하지 못하는 것이다.

디지털 시대에는 여기에 데이터 문맹률도 걱정해야 한다. 데이터를 읽을 수는 있지만 그 데이터의 의미를 이해하지 못한다면 앞으로 데이터 시대를 살아가는 데 많은 어려움이 예상된다. 데이터에 대한 실질 문맹률을 낮추기 위해서는 해결해야 할 문제를 정의하고, 그 문제를 해결하기 위한 데이터를 수집하고, 적합한 방법으로 분석하여, 이를 의사결정에 적극적으로 활용하여 원하는 결과를 만들어 내는 것이다. 데이터 리터러시는 빅데이터에도 있고 스몰데이터에도 있다. 데이터의 종류에 관계없이 문제를 해결할 수 있는 데이터이면 충분하다. 우선은 작은 문제에서 시작하여 데이터로 통찰하는 능력을 키우고 서서히 좀 더 복잡한 문제 해결을 위한 접근 방법을 추구한다면 산업 전반에 걸쳐 데이터를 기반으로 하는 비즈니스가 구현될 것이다. 데이터 리터러시 즉 데이터 문해력을 강화하는 것은 디지털 강국으로 가는 첫걸음이자 생존을 위한 첫걸음이다.



빅데이터 플랫폼 시대... 데이터 연계체계를 확보하라



왜 데이터는 돈보다 더 가치 있는 자산일까

2016년도 미국의 메사추세츠공과대학(MIT)에서 발표한 「데이터 자본의 부상」이라는 기술보고서를 보면 데이터를 4차산업의 신 자본으로 소개하고 있습니다. 우리가 사용하는 돈보다도 더 가치가 있다는 것입니다. 그 이유는 요즘 기업의 가치를 판단할 때 데이터와 같은 무형자산이 매출이나 부동산과 같은 유동자산보다도 더 가치를 인정받기 때문입니다.

또한, 여러 기업에서 동시에 같은 데이터를 비경쟁적으로 활용 가능하며 데이터마다 특성이 있어 비대체적이고, 어떻게 분석하고 활용하느냐에 따라서 무한한 가치를 이끌어 낼 수 있는 경험적 재화의 특징을 가지기 때문입니다.

하나의 예로, 우리나라에서 위암을 진단하는 AI 서비스를 개발한 벤처기업이 우수한 사업 아이템으로 금융 투자는 쉽게 받을 수 있었으나 실상, AI 분석 알고리즘을 개발하기 위해 필요한 의료데이터를 구할 수가 없어 많은 어려움을 거쳐 병원에서 연구용 데이터를 얻어 활용했다는 일화가 있습니다. 이렇게 데이터 확보가 어려운 스타트업이나 중소기업에게는 돈보다도 더 가치 있는 자원임에는 틀림이 없는 것 같습니다.

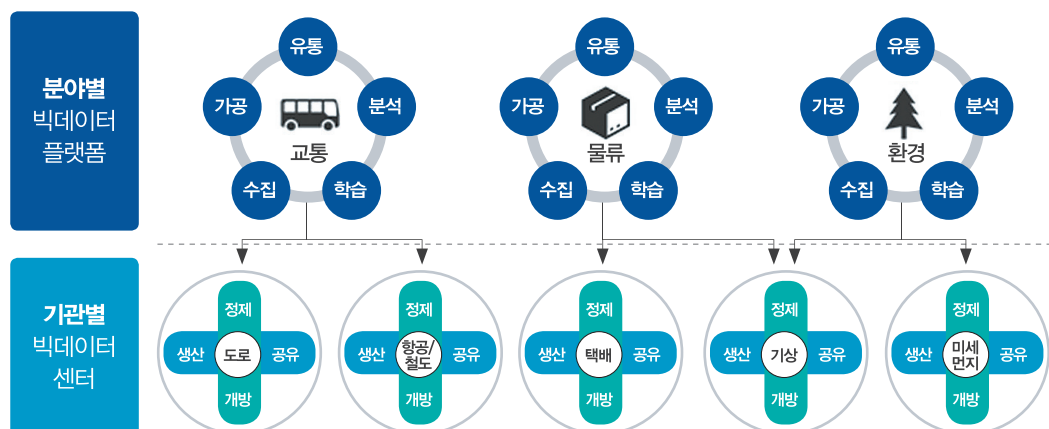
데이터 확산을 위한 빅데이터 플랫폼 및 네트워크 구축 사업 추진

과학기술정보통신부와 한국정보화진흥원에서는 쓸만한 데이터가 넘쳐나고 데이터를 가장 잘 활용하는 나라를 만들고자 하는 목표로 각종 데이터가 모이는 빅데이터 플랫폼 구축 및 데이터 유통·거래 기반 조성을 통해 기존 산업 발전과 신시장 창출 지원에 나서고 있습니다.

이러한 사업을 통해 빅데이터 플랫폼은 10개, 빅데이터 센터는 100개를 구축할 계획입니다. 공공과 민간이 협업하여 활용도 높은 양질의 데이터를 센터에서 생산·구축하고, 플랫폼을 통해 이를 수집·분석·유통·활용 지원할 계획입니다.

현재 빅데이터 플랫폼은 금융, 환경, 문화·미디어, 교통, 헬스케어, 유통, 통신, 산림, 지역경제, 중소기업 경영혁신 등 10개 분야가 선정되어 구축 중이며 각 플랫폼마다 10개의 빅데이터 센터를 운영하게 됩니다.

그림 1 | 빅데이터 플랫폼 및 센터 개념도



기관, 중소기업, 대학 등이 참여하여 빅데이터를 체계적으로 생산·관리하며, 빅데이터 네트워크 조성을 통하여 민·관 협력을 통해 데이터 유통 활용 기반을 조성하고, 플랫폼 간 연계 및 이용 활성화를 지원할 예정입니다.

이 사업은 3개년 사업으로 추진되며 올해 1차년도에는 플랫폼·센터 간 연계·구축을 완료하고, 데이터 품질·표준화 및 유통 가이드라인 마련 등 데이터 생태계 기반을 조성하였습니다. 2차년도에는 분야별 플랫폼 간 연계하고, 플랫폼·센터 간 융합데이터를 재생산하여 시장에 본격 공급 및 혁신서비스 발굴·확산하며, 3차년도에는 중기부 등 범부처 플랫폼과 연계를 확대하고, 분야별 플랫폼 간 융합데이터 생산 등 산업적 활용 확대 및 유통 활성화를 지원할 계획입니다.



핵심 과제를 통한 데이터 생태계 혁신

올해부터 향후 3년간 구축될 빅데이터 플랫폼과 센터는 우리나라의 낙후된 데이터 생태계를 혁신하고 기업의 경쟁력을 확보하기 위한 핵심과제를 추진합니다.

첫째, 양적·질적 데이터 생산·구축 기반 마련

분야별 데이터 부족 문제를 해소하기 위해, 금년 말까지 시장 수요가 높은 1,400여 종의 신규 데이터를 생산·구축하고, 사업이 완료되는 21년까지 총 5,100여 종의 양질의 풍부한 데이터를 생산·구축하여 시장에 공급할 계획입니다.

둘째, 데이터 개방·유통 활성화 기반 조성

기업들이 현행화된 데이터를 확보하는 데도 수개월이 소요된다는 문제점을 개선하기 위해, 센터와 플랫폼 간에는 민간 클라우드 기반으로 데이터 연계 체계를 마련하고, 센터에서 생산된 데이터는 실시간 또는 주기적으로 플랫폼을 통해 개방·공유하여 데이터의 최신성과 연속성을 확보할 계획입니다.

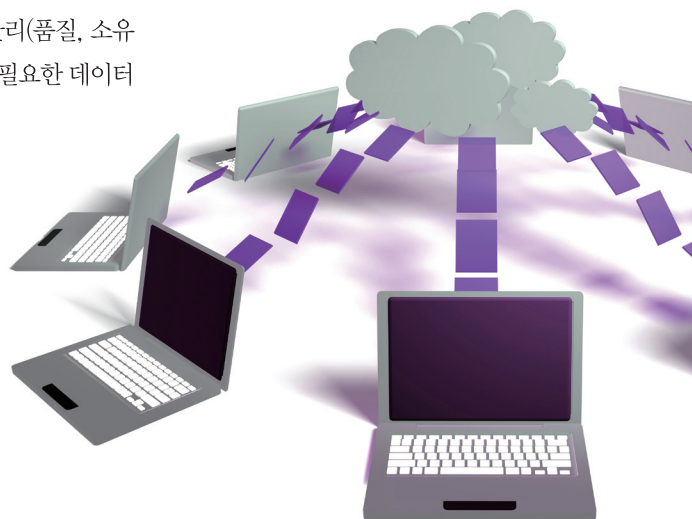
또한, 기존에는 개념이 없었던 데이터를 상품화해 제공하여 데이터 유통거래 시장을 만들어 내기 위하여 데이터 등록 및 감사(품질 등), 검색(소재정보), 사후관리(품질, 소유권 등), 데이터 가격 산정 기준 및 유통 표준약관 등을 마련하여 기업들이 필요한 데이터를 쉽고 빠르게 확인할 수 있도록 지원할 예정입니다.

셋째, 데이터 활용 촉진을 통한 신서비스 시장 창출

100개 센터에서 수집된 데이터는 분야별 플랫폼에서 융합하여 가공·분석하고, 이용 목적에 맞는 맞춤형 데이터 제작 등 양질의 데이터로 재생산하여 기업들이 필요로 하는 데이터를 원하는 형태로 즉시 활용할 수 있도록 제공할 계획입니다.

아울러 다양한 데이터 분석 도구와 AI 학습 알고리즘 등을 제공하고, 일반인도 데이터를 쉽게 분석하고 시각화할 수 있는 사용자 친화

BIG DATA
PLATFORM





적인 분석 환경도 제공할 예정입니다. 또한, 맞춤형 일자리 수요예측, 미세먼지 종합분석, 상권 및 소비트렌드 분석 등 새로운 공익적, 산업적 서비스를 국민과 기업들이 폭넓게 체험하고 활용할 수 있도록 하여 안전한 데이터 활용에 대한 사회적 수용성을 높여 나갈 계획입니다. 이를 통해 플랫폼에서는 분야별 공공, 민간의 각종 현안 문제를 해결하고 데이터로 돈벌 수 있는 시장 환경을 만들어야 합니다.

넷째, 민·관 협력을 통한 데이터 산업 육성기반 구축

분야별 전문성을 가진 부처와 플랫폼 운영기관 등과 민·관 협의체인

데이터 얼라이언스를 구성·운영하여 데이터 기반의 정책, 표준화, 유통 및 제도 등 다양한 현안 사안을 함께 논의하고 데이터의 안전한 활용을 촉진할 수 있는 방안을 마련할 계획입니다.



빅데이터 플랫폼 연계를 통한 데이터 연계·활용 노력 필요

어느 한 분야의 데이터만 분석하는 것보다는 이종 데이터를 결합·분석할 때 예상치 못한 새로운 가치를 발견할 가능성이 높습니다. 예를 들어 조류독감 확산 예측이나 심야버스의 노선을 수립할 때 다양한 데이터를 통해 이러한 문제를 해결합니다.

- 농가 데이터 × 차량이동 데이터 → 조류독감 확산 예측
- 심야시간 통화 데이터 × 택시 승하차 데이터 → 심야버스 노선 수립



이렇게 데이터를 잘 활용하기 위해서는 연계·융합이 무엇보다 중요하지만 정보통신 기술(ICT)이 다양한 현장에서 활용되는 현실에서 데이터가 어디에 있는지 모르거나, 또는 데이터 확보에 많은 시간·비용이 소요되는 어려움이 있습니다.

따라서 우리도 데이터 간에 연계를 위한 표준화 작업과 함께 플랫폼 간에 데이터 협력 프로젝트 조성, 데이터 공유 활용 등의 노력이 시급합니다. 이를 위하여 빅데이터 플랫폼 연계를 통한 데이터 연계·활용 확산을 위해 아래와 같은 노력이 진행 중에 있습니다.

첫째, 한곳에서 데이터를 통합 검색하고 활용 할 수 있도록 플랫폼 간에 생산구축한 데이터 목록을 서로 공유하고 데이터 간에 연결체계를 구축

최근 구글의 데이터셋 서치 베타(<https://toolbox.google.com/datasetsearch>)에서도 DCAT과 schema.org 기반 데이터셋 목록을 수집하고 통합 검색할 수 있도록 지원하고 있습니다. 우리나라도 플랫폼 간 데이터 연계를 위하여 다년간 표준화로 메타데이터 서술, 교환 기능과 데이터 유통/품질 등의 다양한 정보 공유가 가능한 DCAT v2와 schema.org에서 제공하는 풍부한 어휘사전 등을 채택하여 플랫폼 간에 연동을 통하여 생산한 모든 데이터가 한곳에서 통합검색 가능하도록 데이터 목록을 상호 공유할 계획입니다.

올해에는 빅데이터 플랫폼에서 제공하는 데이터 목록을 데이터 통합지원센터에서 연계 수집하여 통합 검색 지원하는 시맨틱 데이터맵을 제공할 계획입니다. 이용자들은 통합 검색을 통하여 데이터 위치, 소유권, 설명서, 품질, 유통, 활용 등 각종 데이터 속성 정보를 확인하고 이용 가능하며 시맨틱 검색을 통하여 데이터셋 간 유사도에 따른 추천 데이터도 이용할 수 있습니다.

둘째, 개방된 데이터를 보다 쉽게 연계 활용 할 수 있도록 데이터 개방 포맷은 기계 판독이 가능한 형태의 오픈 포맷 방식(CSV, RDF, LOD 등)을 적용

오픈 포맷이란 특정 소프트웨어(아래이한글·엑셀 등)에 종속되지 않고 다

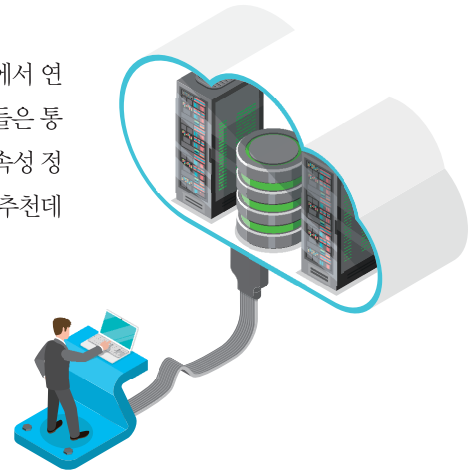
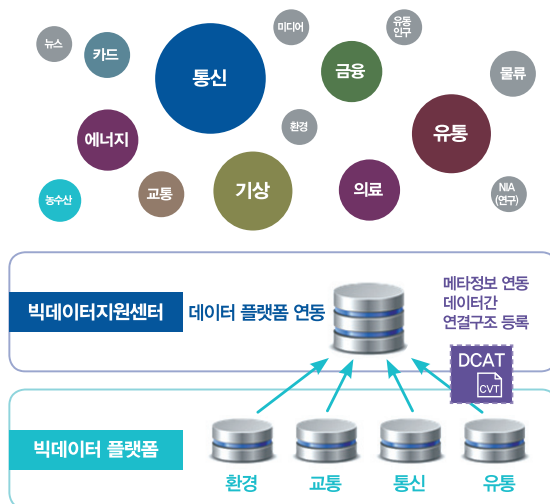


그림 2 | 시맨틱 데이터 맵 개념도



• 데이터맵 기반 이용환경 조성

(LOD, CSV 등 공유데이터 활용)

- LOD : Linked Open Data로 개방형 데이터 표준
- CSV : Coma Separated Value로 침표를 기준으로 항목구분 텍스트 데이터

• 개방/유통 데이터 메타정보 통합 관리

• 메타데이터 수집 및 연동(DCAT)

(개방/유통 데이터 메타정보 업데이트)

- DCAT : Data Catalog Vocabulary로 웹기반 데이터 카탈로그의 상호연동 표준

• 오픈 데이터 플랫폼 구축 운영

(메타정보 표준 규격화)

양한 소프트웨어에서 자유롭게 활용할 수 있는 형태의 데이터를 말합니다. 현재 국내외 공공데이터 개방에 활용 중인 텀버너스리가 제안한 5개 단계 포맷을 참조하여 3~5단계에 해당하는 데이터 포맷을 적용하여 데이터 개방을 추진할 예정입니다.

그림 3 | 데이터 개방 포맷

구분	1단계	2단계	3단계	4단계	5단계
기계판독이 가능한 형태	미충족포맷 (포털등록불가)	최소충족포맷	오픈포맷		
특징	특정 소프트웨어에서 읽을 수만 있는 데이터로 자유로운 수정, 변환 불가	특정 소프트웨어에서 읽고 수정, 변환 가능	모든 소프트웨어에서 읽고 수정, 변환 가능	통합 자원 식별자(URI)를 기반으로 데이터 속성 특성 관계를 기술하고 있는 데이터 구조	웹상의 다른 데이터와 연결, 공유 가능
예시	PDF	HWP, XLS, JPG, PNG, WMV, MPEG, MP3, SWF	CSV, JSON, XML	RDF	LOD

셋째, 데이터 구축과 활용 관련 표준용어 및 포맷 활용을 증대

빅데이터 플랫폼에서 데이터를 구축하고 개방하는 데 있어 용어의 정의가 없어 의미의 경계가 명확하지 않아 사용자의 자의적 해석 및 정보사용의 오류가 발생할 수 있습니다. 또한, 이음동언어 등의 사용으로 데이터 구조 간 중복, 불일치로 인하여 기관 및 사용자 상호 간의 의사소통이 원활하지 못한 문제가 존재합니다.

따라서, 공공과 민간에서 개별적으로 관리되고 있는 데이터의 용어, 형식, 항목 등을 단계적으로 표준화하여 데이터 생산·구축 및 연계활용에 적용하고자 합니다. 현재, 공공분야에서는 행안부에서 제공 중인 행정표준용어(2019.10 현재 약 1만 4천 개, 행정표준용어.xls 참조)가 활용 중이며 행정표준코드관리시스템(<http://www.code.go.kr>)을 통하여 확인 가능합니다.

넷째, 데이터 연계키 활용을 활성화

플랫폼 및 센터에서 수집된 데이터를 플랫폼 간 또는 센터들과 연계하기 위해 데이터 연계키가 반드시 필요하고 표준화가 되어야 합니다. 이를 통하여 서로 다른 분야별 다양한 데이터를 연계 융합하여 빅데이터 서비스 확산이 가능합니다. 다양한 분야에서 공통으로 활용할 표준 연계키에 대한 발굴과 관련 표준 포맷의 지정 필요하여 그간 활용이 많았던 ‘사람’, ‘기업’, ‘공간’ 및 ‘시간’ 영역으로 구분하여 표준 연계키를 우선 발굴하여 활용할 예정입니다.



그림 4 | 표준 연계키 발굴 범위

영역	표준 연계키
사람	성, 연령, 연령대, 이메일, 거주지 주소, 우편번호
기업	사업체 명칭, 사업체 전화번호, 사업체 주소, 법인번호, 업종
시간	시간, 시각, 날짜, 연도
공간	주소, 좌표, 행정구역, 국가기초구역, 건물 (집합건물, 빌딩)

예로서, ‘공간’에 대한 표준 연계키는 지역을 식별할 수 있는 주소, 좌표, 행정구역 항목 등이 이용되며 이를 활용하여 아래와 같이 이중데이터의 융합이 가능합니다.

그림 5 | ‘행정구역’ 연계키를 이용한 이중 데이터의 융합

여성인구	남성인구	동 명칭	미세먼지	초미세먼지	동 명칭
15,971명	13,432명	강남구 논현1동	10.5	3.4	강남구 논현1동
27,318명	25,848명	강남구 논현2동	11.4	3.9	강남구 논현2동
:	:	:	:	:	:

표준연계키 : 행정구역

‘시간’에 대한 표준 연계키는 시간을 구분하는 연도, 연/월/일 등을 표준 연계키로 활용 가능합니다.

그림 6 | '연도' 연계키를 이용한 경제 데이터의 결합

고용증가율	연도	성장률	연도
5.6%	2013	3.4%	2013
7.8%	2014	4.4%	2014
:	:	:	:

표준연계키 : 연도

이제 빅데이터 클라우드의 활성화를 준비해야 할 중요한 시점

빅데이터, 인공지능(AI), 5G서비스 등의 활성화를 위해서는 원스톱으로 데이터를 분석하고 학습하며 서비스로 제공할 수 있도록 탄실한 기반 인프라로 빅데이터 클라우드 서비스의 활성화가 필요합니다. 빅데이터 클라우드 서비스는 기존 데이터 저장 개념에서 확장되어 데이터 수집·가공·분석·시각화는 물론 AI학습, 서비스 제공까지 클라우드 서비스로 제공하는 개념입니다.

그동안 빅데이터 기술은 매우 빠르게 발전해 왔습니다. 요즘은 센서, 이미지, 영상 등 다양한 비정형 데이터를 실시간 분석하고 딥러닝, 시각화하는 기술까지 오픈소스 방식으로 너무나도 잘 개발되어 배포되고 있습니다. 외국에서는 아마존, 마이크로소프트사 등 우수기업들이 이러한 빅데이터 분석, 딥러닝, 시각화 기술들을 클라우드 서비스에 접목시켜 빅데이터 클라우드 서비스로 제공하여 누구나 쉽게 데이터를 다루고 수준 높은 빅데이터·AI서비스를 만들어 낼 수 있는 환경을 제공하고 있습니다.

하지만 우리나라는 이러한 기술들을 활용하는 능력이 부족한 상황입니다. 따라서, 기존의 저장소 개념의 클라우드에서 오픈소스로 제공 중인 우수한 빅데이터 기술들을 탑재한 빅데이터 클라우드 서비스를 플랫폼에서 활용하는 것이 시급합니다. 앞으로 플랫폼에서도 빅데이터 클라우드를 잘 갖춘다면 성능 좋은 빅데이터 서비스를 신속하게 개발해 제공할 수 있을 것입니다.

현재 빅데이터 클라우드 서비스에 활용될 수 있는 국내 오픈 데이터 플랫폼으로는 한국 전자통신연구원에서 국책과제¹⁾로 개발 중인 SODAS(Smart Open Data As a Service) 플랫폼이 있습니다. SODAS는 국제 표준인 DCAT를 활용해 수많은 기관 간에 데이터셋을 상호 공유하며 클라우드 기반으로 데이터를 수집·분석·활용하고 품질 및 유통 등을 관리할 수 있도록 아키텍처가 구현되어 국내 플랫폼 활성화에 활용 및 참고할 만합니다.

우리나라가 핵심기술은 미흡하지만 2000년대 통신 인프라를 잘 만들어 초고속인터넷, BcN, 스마트폰 시대를 거쳐 ICT의 많은 성과와 발전을 만들어 냈듯이 이제 좋은 데이터로 다양한 혁신 서비스를 만들어 낼 수 있도록 빅데이터 클라우드의 활성화를 준비해야 할 중요한 시점이 되었습니다.

¹⁾ 국제표준 기반 오픈 데이터 유통 플랫폼 확장 개발('17.03 ~ '19.12, 주관기관 : 한국전자통신연구원) : 방대한 오픈 데이터의 연계·활용성과 품질 수준 향상을 위하여 국제표준을 준용하는 데이터 관리 인프라, 상호연동 및 데이터 서비스 기술 연구 개발.



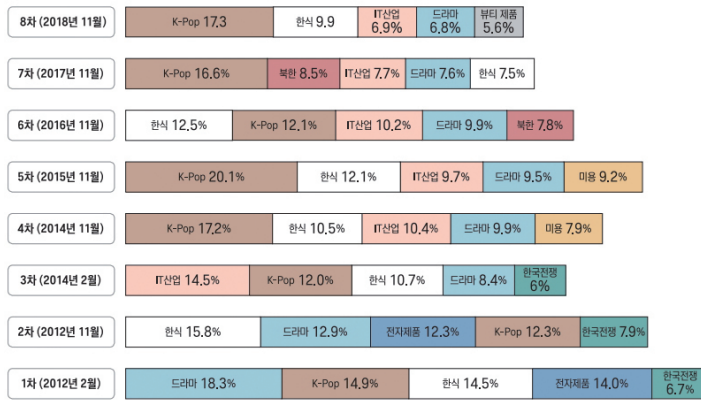
한류 콘텐츠의 인기 비결을 알고 싶다면 통계를 찾아보자

얼마 전까지만 하더라도 ‘코리아(KOREA)’라고 하면 ‘동아시아에 있는 분단 국가’ 또는 ‘과거 6.25 전쟁을 겪었던 가난한 나라’ 쪽으로 인식하거나, 이마저도 모르는 사람들은 대한민국과 북한을 혼동하는 경우도 많았다.

그랬던 KOREA가 최근 들어서는 과거와 전혀 다른 문화강국의 이미지로 전 세계에 각인되고 있다. 바로 ‘한류(韓流)’라는 콘텐츠 덕분이다. ‘K-팝(K-POP)’이 이끄는 음악을 중심으로 드라마와 음식, 화장법과 같은 다양한 콘텐츠들이 대한민국의 이미지를 새롭게 바꾸고 있다.

그렇다면 이들 한류 콘텐츠는 얼마나 다양한 분야에서 대한민국의 이미지를 바꾸고 있을까. 우리의 음악이 얼마나 아름다운지, 우리의 드라마가 어째서 재미있는지에 대한 이유를 단순히 '좋다'라는 표현이 아니라 통계를 통해 보다 명확하게 알아보자.

한국을 연상시키는 이미지



(출처 : 한국콘텐츠진흥원(KOCCA))

한류의 탄생과 진화

한류는 우리의 문화 콘텐츠가 해외에서 인기리에 소비되는 현상을 가리키는 신조어다. 신조어인 만큼 탄생 배경에 대해서는 학자마다 조금씩 다른 주장을 펼치고 있지만, 대체적으로 1990년대 후반부터 본격적으로 사용되기 시작했다는 의견이 지배적이다.

이에 대해 한국국제문화교류진흥원(KOFICE)의 관계자는 “한류라는 신조어는 중국에서 출발한 것으로 알려져 있다”라고 언급하면서 “특히 지난 1997년에 인기리에 방영되었던 국내 드라마가 중국에서 폭발적인 인기를 누리면서 한류라는 단어가 탄생했다”라고 말했다. 한국국제문화교류진흥원은 한류 확대를 주요 업무로 하는 문화관광부 산하의 민간 재단이다.

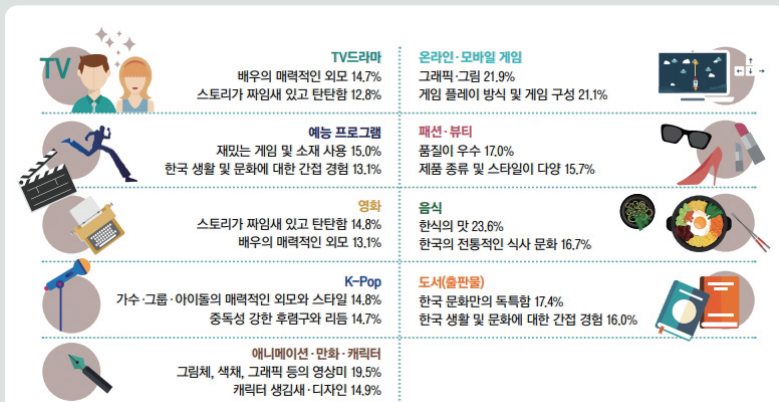


이처럼 한류는 드라마를 통해 탄생했지만 만약 드라마 위주의 수출에만 머물렀다면 오늘날의 한류는 존재하지 않았으리라는 것이 전문가들의 공통된 생각이다.

드라마의 인기가 한풀 꺾일 무렵인 당시의 한류는 장르

와 대상 연령, 그리고 지역적 한계에 직면하면서 새로운 동력이 요구되던 시기였다. 그 때 등장한 것이 바로 한류의 세계화를 이끈 'K-팝'으로서, 업계에서는 이 시기를 초기 한류와 구분하여 '한류 2.0'이라고 부른다.

한류 콘텐츠 유형별 인기요인



(출처 : 한국국제문화교류진흥원(KOFICE))

‘한류 1.0’이 드라마를 통해 중국에서 유행하기 시작했다면, ‘한류 2.0’은 K-팝이 인기를 끌기 시작한 일본을 무대로 알려지기 시작했다. 현지화라는 전략을 통해 일본에서 인기를 끌기 시작한 K-팝은 국내 아이돌 그룹의 노래들이 대거 일본의 가요 차트를 석권하면서 가능성을 확인했다.

중국에서 드라마로 한류가 탄생했고 일본에서 음악으로 한류가 꽃을 피울 수 있었던 이유에 대해 한국국제문화교류진흥원의 관계자는 “무엇보다 우리의 콘텐츠가 우수했기 때문이다”라고 강조하면서도 “국가별 진출 시기가 당시 상황과 맞물려 적절하게 이루어진 점도 원인 중 하나”라고 밝혔다.

실제로 국내 드라마가 인기를 끌기 시작한 1990년대의 중국은 자본주의의 물결이 들어가면서 가족 해체 현상이 나타나던 시기였다. 그리고 K-팝이 본격적으로 알려지기 시작한 2000년대 초반의 일본은 음반 산업이 침체기에 머물던 시기였다.

이처럼 K-팝은 침체된 일본 음반 시장의 돌파구가 되었지만, 그 범위는 대부분 아시아 시장에 머물렀다. 팝의 원조 국가들이라 할 수 있는 유럽이나 미국에서는 활동이라는 말조차 하기 어려울 정도로 그 성과가 미미했다.

K-팝으로 한류 2.0 시작

아시아권에서만 인기를 끌었던 K-팝은 2012년에 접어들면서 전혀 예상치 못한 노래로 재평가를 받게 되었다. 바로 싸이가 부른 '강남스타일'이다. 이 노래의 뮤직비디오는 동영상 공유 사이트인 유튜브의 '세계에서 가장 많이 본 동영상' 순위에서 1위를 달릴 정도로 엄청난 인기를 끌었다.

팝송과 관련된 순위 중 세계 최고의 권위를 자랑하는 미국 빌보드에서 2위까지 오르는 기염을 토했고 전 세계 대부분의 사람이 뮤직비디오에 등장하는 '말춤'을 따라하는 등, 강남스타일은 국내 음악 역사의 새로운 획을 그었다.

이 같은 결과에 대해 한국국제문화교류진흥원의 관계자는 "강남스타일의 폭발적 인기는 K-팝도 세계에서 통할 수 있다는 성공사례를 보여줬다"라고 평가하며 "이 노래 하나가 한류를 전 세계로 전파시킨 원동력이 되었음을 부인할 수 없다"라고 덧붙였다.

하지만 '강남스타일'은 치밀한 기획으로 성공한 것이 아니라 일종의 '로또' 같은 행운이 따른 노래였다. K-팝을 전 세계가 인정해주는 음악의 주류(主流)로 편입시켰다고 보기는 어려웠던 것이다. 따라서 K-팝이 본격적으로 성공하려면 단발성의 노래가 아닌 K-팝을 이끌어 나갈 진정한 스타가 필요했다.

아이돌 그룹인 '방탄소년단(BTS)'이 등장한 것은 바로 이 무렵이다. 이들은 단발성이 아닌 다양한 노래를 통해 K-팝을 전 세계에 알렸다. 놀라운 점은 BTS가 음악에 관한 가장 빠르게 인기를 얻을 수 있는 무대인 미국 진출을 처음부터 목표로 하지 않았다는 점이다.

그 대신 BTS는 유튜브를 통해 전 세계에 자신들의 노래를 들려주고, 소셜미디어를 활용하여 해외 팬들과 직접 소통하는 방식을 택했다. 그 결과, 이들은 빌보드를 포함하여 전 세계에서 내로라하는 음악 차트를 석권하면서 세계적인 아이돌 그룹으로 성장할 수 있었다.

그 결과 방탄소년단은 미국 빌보드 차트를 연이어 석권하며 명실상부한 글로벌 가수로 자리를 잡았다. 또한 걸그룹으로 정상의 인기를 누리고 있는 트와이스는 일본 데뷔 앨범을 25만 장이나 판매했을 뿐만 아니라 올해 초 발매한 일본 싱글 앨범으로 오리콘 차트 1위 자리를 차지하기도 했다.



빌보드 1위를 차지한 BTS. (출처 : 빌보드 홈페이지)

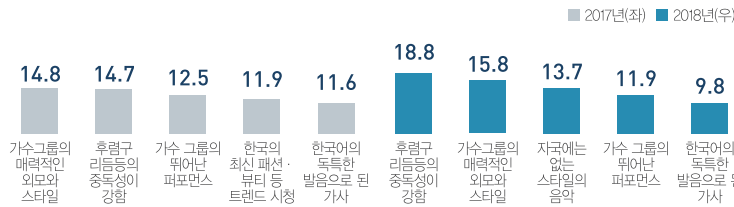
한류 콘텐츠의 발전으로 신 한류도 기대

BTS를 중심으로 여러 그룹이 이끄는 K-팝이 한류의 중심이라는 사실은 한국국제문화교류진흥원이 지난해 발간한 '2018 해외한류실태조사' 보고서에서도 확인할 수 있다. 해당 보고서에 따르면, 지난 2017년부터 한류 콘텐츠의 인기 및 소비는 K-팝이 주도했다고 해도 과언이 아닌 것으로 나타났다.

보고서에 제시된 통계 수치를 살펴보면 한국에 대한 연상 이미지로 K-Pop을 떠올리는 한류콘텐츠 소비자가 16.6%로 가장 많았고, 한국을 연상시키는 제품에서는 지난 2016년 7위였던 K-Pop이 2년 만에 3위까지 상승한 것으로 드러났다.

또한 한류콘텐츠 호감도 부문에서도 '한국 K-팝 콘텐츠가 마음에 든다'라는 응답이 2016년에 비해 18.2%나 증가하여, 전년에 대비해 20.1%가 증가한 예능프로그램 다음으로 증가 폭이 높은 것으로 파악됐다.

K-팝 인기요인



(출처 : 한국콘텐츠진흥원(KOCCA))

이 외에도 한류콘텐츠 인기도 부문에서는 한식과 패션·뷰티 다음으로 38.9%를 기록한 K-팝이 인기 있는 것으로 나타났다. 또한 전년에 대비하여 최고로 많이 증가한 5.7%의 애니메이션에 이어, 전년 대비 증가 값이 두 번째로 높은 5.5%를 기록했다.

'자국에서 인기 있는 한국 제품'이라는 응답에서도 역시 2016년에 대비해 K-팝의 인기가 11% 이상 큰 폭으로 상승하여 47.1%를 기록했으며, 만나고 싶은 한류스타에 대한 조사 결과에서도 K-팝의 강세가 반영된 결과가 나타났다.

이 같은 통계치는 국내에서 인기를 구가하고 있는 K-팝 아이돌들이 해외 시장에서 성공적으로 정착할 수 있는지를 판단하는 가늠자 역할을 할 수 있다는 것이 전문가들의 평가다.



이 외에도 K팝의 글로벌 돌풍 현상은 한국에 대한 연상 이미지나 자국에서 인기 있는 한국 제품에 대한 견해뿐 만 아니라 한류 콘텐츠의 소비나 향후 전망 등에서도 찾아볼 수 있다.

한류 콘텐츠의 소비 비중은 2016년에 비해 전체적으로 증가했는데, 특히 K-팝의 소비 비중 증가율이 9.1%로 모든 조사 콘텐츠 항목 중 가장 많이 상승한 것으로 나타났다.

향후 한류 콘텐츠와 관련된 활동 참여 의향 역시 K-팝 동호회에 대한 참여 의향이 지난해에 비해 11.3%나 늘어나서, 작년과 대비하여 증가폭

이 가장 큰 활동으로 조사되었다.

이와 관련하여 한국국제문화교류진흥원의 관계자는 “더 이상 한류는 주변 국가나 일부 콘텐츠에만 해당하지 않고 새로운 유통 플랫폼을 찾아 또 다른 도약을 준비하고 있다”라고 언급하며 “아시아권을 대상으로 드라마나 음악 등이 이끈 한류가 이전의 모습이었다면 이제는 우리나라의 모든 문화 앞에 ‘K’자를 붙이고 있다는 것이 가장 큰 특징”이라고 말했다.

이를테면 우리의 화장법을 알려주는 K-뷰티나 음식을 맛보게 해주는 K-푸드, 또는 우리의 문화를 체험하게 해주는 K-컬처 등과 같은 신조어를 통틀어 ‘한류 3.0’이라고도 부르는데, 이제는 여기서 한발 더 나아가 ‘신(新) 한류’의 등장도 기대해 볼만하다는 것이다.

키즈 콘텐츠의 등장과 플랫폼의 변화

한편 한류콘텐츠는 최근 들어 새로운 국면을 맞고 있다. 바로 ‘키즈(kids)’ 콘텐츠다. 최근 국내 키즈 산업을 들여다보면 영상 콘텐츠가 대부분의 주도권을 잡고 있는 형국이다.

키즈 콘텐츠가 지닌 영향력은 유튜브 채널의 구독자 순위를 살펴보면 보다 명확해진다. 불과 얼마 전까지만 해도 유튜브 채널의 대표 콘텐츠였던 뷰티나 먹방 관련 동영상들보다 어린이들의 일상과 놀이를 주요 내용으로 하는 ‘키즈 콘텐츠’가 더 높은 자리에 위치해 있음을 알 수 있다.

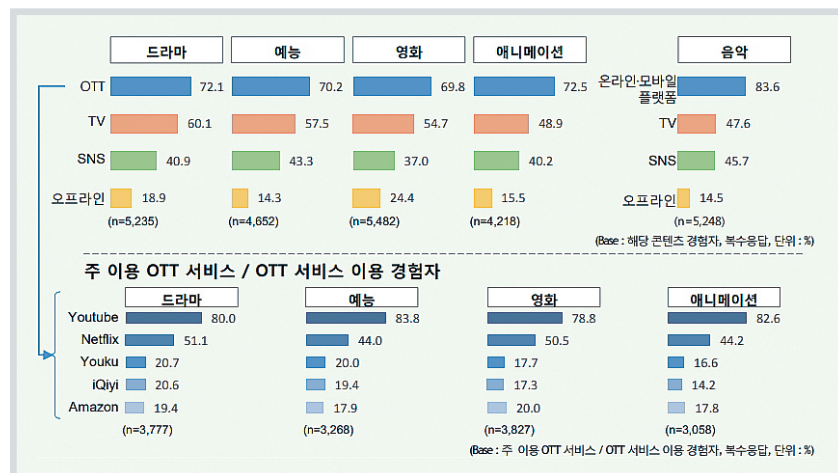
유튜브 분석 통계 전문 사이트인 소셜블레이드의 분석 결과를 살펴보면, 구독자를 기준으로 국내 상위 13개 채널 중 8개가 키즈 콘텐츠를 제공하는 채널인 것으로 나타났다. 이들 채널의 구독자 수를 모두 더하면 무려 3,500만 명에 달한다.

중요한 점은 이 같은 추세가 비단 우리나라만의 현상이 아니라는 점이다. 해외에서도 유튜브나 넷플릭스 같은 글로벌 플랫폼을 중심으로 키즈 콘텐츠 소비가 활발하게 진행되고 있는 양상이다.

실제로 미 경제지인 포브스(Forbes)가 발표한 '세계 최고 수입의 유튜브 스타 2018'에서 1위를 차지한 것은 장난감 체험 채널을 운영하는 7살 소년인 것으로 드러났다.

키즈 콘텐츠와 함께 주목받고 있는 한류 콘텐츠의 또 다른 특징이라면 스트리밍 서비스를 통한 콘텐츠의 이용 증가 경향을 들 수 있다. 기술 발전으로 네트워크 속도가 개선됨에 따라 기존 다운로드 방식을 취했던 콘텐츠 이용이 스트리밍(streaming)으로 전환되고 있는 것.

음원에서부터 출발하여 현재 동영상 서비스에도 스트리밍 방식이 적용되고 있고, 스트리밍 이용자가 늘어남에 따라 한류콘텐츠들도 온라인 또는 모바일 스트리밍 서비스를 통해 유통되고 있는 추세다.



변화하고 있는 한류 콘텐츠 접촉 경로 (출처: 한국콘텐츠진흥원(KOCCA))

실제로 한류 콘텐츠와 관련된 상품을 소비하는 데 있어 온라인 스트리밍 이용이 얼마나 증가하고 있는지를 파악하기 위해 지난 2017년부터 해외 한류 실태조사에서 스트리밍과 다운로드를 나누어 한류콘텐츠 이용 경로에 대한 조사를 실시하고 있다.



그 결과 온라인 또는 모바일 스트리밍을 통한 한류콘텐츠 이용이 현저히 늘어난 것으로 조사되었다. 먼저 영상 콘텐츠 영역을 살펴보면, 2017년에는 2016년에 비해 TV 이용은 감소한 반면에 대다수의 콘텐츠 분야에서 무료 온라인·모바일 스트리밍 이용이 TV를 통한 이용을 앞지른 것으로 나타났다.

예능의 경우, 작년에는 TV를 통해 시청하는 이용자가 가장 많았던데 반해 2017년에는 무료 스트리밍을 통한 이용자가 가장 많았다. 2017년 무료 온라인·모바일 스트리밍 이용 비율은 62.6%로, 한국 예능프로그램을 접하는 경로 중 가장 큰 비중을 차지했다. 반면 TV 이용은 2016년 대비 0.4%가 감소하여 60.2%에 그쳤다.

애니메이션의 경우도 비슷하다. 2017년 들어 온라인과 모바일의 스트리밍 이용율은 60.8%로서 TV 이용률인 53.5%를 앞지른 것으로 나타났다.

애니메이션의 경우도 비슷하다. 2017년 들어 온라인과 모바일의 스트리밍 이용율은 60.8%로서 TV 이용률인 53.5%를 앞지른 것으로 나타났다.

TV드라마의 경우에는 TV를 통한 시청이 2016년에 비해 0.7% 줄어들었고, 영화는 TV 이용이 전년에 비해 6.1% 정도 증가했다. 하지만 온라인과 모바일의 스트리밍 이용은 11.8% 정도 증가하여 TV 이용 비율을 온라인과 모바일 스트리밍 이용 비율이 많이 따라잡은 것으로 나타났다.

K-팝 분야 역시 온라인과 모바일상의 스트리밍을 통해 음악을 접하는 비중은 62.3%로서 TV 이용 비중인 55%를 앞질러 가장 높은 것으로 나타났다. 이들 간의 차이는 7.3%로서, 영상 콘텐츠 중 가장 그 차이가 큰 쪽으로 나타난 애니메이션 분야와 동일했다.

이렇게 몇년 사이 한류 콘텐츠를 온라인과 모바일 스트리밍을 통해 이용하는 이용자들이 급증했다는 조사 결과는, 앞으로 온라인 그리고 모바일을 통한 스트리밍이 전 세계적인 주요 한류콘텐츠 유통경로가 될 것이라는 사실을 짐작하게 만든다.

또한 앞으로 한류 콘텐츠를 확산시키기 위해서는 스트리밍 플랫폼 활용 전략이나 다양한 스트리밍 관련 정책 마련이 중요해졌다는 시사점도 제공한다.

이러한 발전의 지속세를 이어가기 위해서는 국수주의적인 시각에서 벗어나 세계와 쌍방향으로 소통하면서도 기술의 흐름을 읽는 적극적인 자세가 필요하다는 것이 전문가들의 공통된 의견이다.

최근에는 4차 산업혁명의 기대감과 함께 블록체인과 콘텐츠의 결합으로 인한 활용성에 대한 업계의 관심이 증가하고 있다. 따라서 다양한 장르와의 결합 등을 통해 융합적 생태계가 더 빠르게 확산될 수 있을 것으로 점쳐지고 있다.

과거의 성공에 안주하지 않고 더 나은 미래를 바라보며 현재를 융합하려는 노력, 이것이 이야말로 한류의 성공을 이어가기 위해 지금 필요한 자세일 것으로 보인다.

지역 자료가 궁금하다면... 「e-지방지표」클릭!

e-지방지표에서 다양한 시도·시군구 자료를 찾아보자

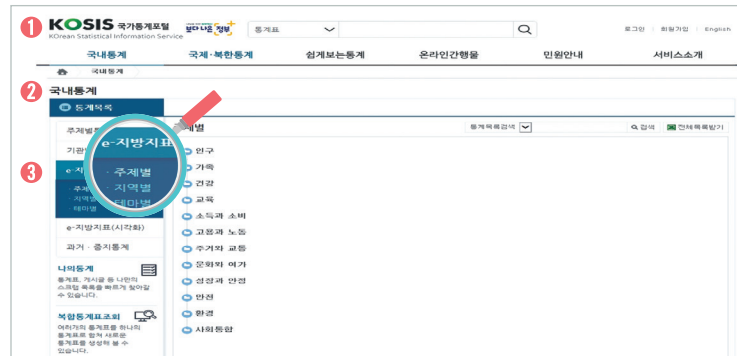
통계청에서는 KOSIS, SGIS, MDIS 등 다양한 방법으로 통계자료를 제공하고 있다. 그 중 KOSIS 내에 e-지방지표는 시도/시군구 자료를 한 곳에서 확인할 수 있어 국가균형 발전과 자치분권이 강조되고 있는 현 시대에 꼭 필요하다고 볼 수 있다.

..



e-지방지표 찾아가기

① <http://kosis.kr> → ② 국내통계 → ③ e-지방지표(통계표, 시각화)



e-지방지표 통계표 서비스

e-지방지표 시각화 서비스



e-지방지표는 12개 분야 222개 지표*를 포함하고 있으며 통계표 서비스와 시각화 서비스를 하고 있다. 통계표 서비스는 주제별, 지역별, 테마별로 분류하여 서비스하고 있으며, 시각화 서비스는 지역간 비교자료를 지도와 그래프 등을 통해 보여준다.

* 지표 구성

	인구	가족	건강	교육	소득과 소비	고용과 노동	주거와 교통	문화와 여가	성장과 안정	안전	환경	사회 통합
222	21	10	14	14	14	23	27	10	33	18	16	22

여기서는 e-지방지표를 활용해 지역별 자료를 살펴보고자 한다.

자료 이용하기

- 가급적 e-지방지표에 수록된 자료 위주로 작성했으며, 일부 자료는 작성부처 자료를 참고하여 작성함(출처 표시).
- 최근 자료를 기준으로 10년 전(20년 전)과 비교했으며, 일부 자료는 해당 자료 제공 시작 시점과 비교함.
- 활용한 지표의 이해를 돕기 위해 통계설명자료(메타정보)를 인용하여 설명하였으며, 일부 용어는 네이버에서 인용하여 설명함.

▶ 제주도에 진짜 여자가 많을까?



남녀성비는 주민등록인구를 기준으로 여자 백명당 남자 수로 정의하고 있다. 성비가 100을 초과하면 남자인구가 여자인구보다 많은 것을 의미하고, 반대로 성비가 100미만이면 여자인구가 남자인구보다 많은 것을 의미한다.

1998년 100.8명이었던 우리나라 남녀성비는 2008년 100.4명이었다가 2014년 100.0명을 기록한 후 100명 미만으로 변화되어 2018년 99.6명으로 감소되었다.

지역별로는 어떤 변화가 있었을까? 울산광역시 북구의 남녀성비는 1998년 108.4명으로 여자 100명당 남자의 수가 가장 많은 것으로 나타났고 의령군이 92.5명으로 여자 100명당 남자 수가 가장 적은 것으로 나타났다.

2018년 남녀성비는 용진군이 131.0명으로 여자 100명당 남자의 수가 가장 많은 것으로 나타났고, 경상북도 영덕군은 90.5명으로 여자 100명당 남자 수가 가장 적은 것으로 나타났다.

돌과 바람과 여자가 많다는 제주도는 1998년 98.3명, 2008년 99.8명이었지만 2009년 100.0명을 기록하였고 2018년 101.3명으로 남자가 더 많아진 것으로 나타났다. 혹시 남자가 많은 곳을 찾으신다면, 혹은 여자가 많은 곳을 찾으신다면 KOSIS에서 주제별<인구·가구>주민등록인구현황을 검색해보시라! 내가 원하는 나이의 남녀인구를 확인할 수 있다.



[지표출처 : e-지방지표>남녀성비(시도/시/군/구),
주민등록인구현황>행정구역(시군구)별, 성별 인구수]



▶ 출산율이 가장 높은 지역은 어디일까?

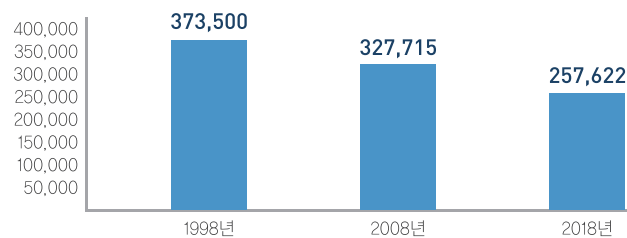


혼인은 사회의 기초적 구성단위인 가정·가족을 형성하는 단서가 되고 모든 사회가 어떤 형태로든 혼인을 승인하고 이에 법적 규제를 하는데 형태는 각 사회의 경제적 종교적 민족적 요소에 따라 다르다.

우리나라에서 혼인은 남편과 아내로 구성된 법적관계, 즉 법률혼주의 원칙에서 혼인 신고수리의 경우와 직권에 의해 가족관계 등록부 기재에 의한 신고로 이뤄진다. 신고에 의한 것이므로 혼전 동거, 별거 등 사실혼과는 차이가 있을 수 있으며 남편의 주소지를 기준으로 행정구역별로 집계한 자료이다.

지역별 혼인건수를 살펴보면 수원시가 1998년 8,315건, 2008년 7,997건, 2018년 6,902건으로 가장 많았으나, 주민등록인구 1,000명당 혼인건수는 1998년 이전시 10.4

혼인건수(건)



혼인건수는 1998년~2018년 20년 사이 115,878건, 31.0%가 감소했다. *(1998년) 373,500 (2008년) 327,715 (2018년) 257,622

건, 2008년에는 오산시와 거제시가 9.3건, 2018년에는 화천군이 7.6건으로 가장 많은 것으로 나타났다.

혼인이 감소한 만큼 출산율도 급격히 감소했다. 임신을 할 수 있는 여성의 숫자가 많을 수록 인구가 늘어날 가능성도 높아진다. 여성 사망자와 여성 출생자수가 비슷해 평형을 이루려면 최소 2.1명의 자녀를 낳아야하는데 이것을 인구대체수준이라 부른다. 자녀수가 인구대체수준에 이르지 못하면 전체 인구가 계속 줄어든다.

우리나라의 현재 출산율은 0.98명에 불과하다. 2000년 1.48명이었던 우리나라 합계 출산율은 2018년 0.98명으로 마카오를 제외하고 세계 유일 1명대 미만 국가이다. 그럼에도 불구하고 해남군의 2018년 합계출산율은 1.886으로 전국 최고를 기록했다. (구단위 : 부산 강서구 1.611, 시단위 : 세종특별자치시 1.566, 도단위 : 전라남도 1.240) 해남 땅끝마을에서 울려 퍼지는 아이 울음소리가 전국 곳곳에서도 들리길 바라본다.

매년 아이 울음소리가 줄어드는 와중에 출산율이 증가한 곳도 있다. 부산광역시 강서구의 경우 2000년 1.219명에서 2018년 1.611명으로 0.392명 증가했다. 부산광역시 강서구는 어떻게 출산율을 높였는지 주의 깊게 검토해봐야 할 일이다.

합계출산율이 낮아지다 보니 가족 구성원의 변화도 있다.

3세대 이상이 같이 모여살거나, 자녀의 수가 많아 가족구성원이 많았던 가족 구성에서 부부와 미혼자녀만으로 이루어진 핵가족의 형태를 보이다가, 최근에는 가구원이 한 명인 1인 가구가 증가하는 추세이다.

1980년 7인 이상의 가족을 이뤘던 가구는 15.2%였으며, 그 중 대구시가 40,414가구로 가장 많았다. 2018년에는 7인 이상 가족 가구는 0.3%로 나타났으며 그 중 용인시가 1,205가구로 가장 많은 것으로 나타났다.

연도	전체가구	7인 이상 가구		1인가구	
			비율		비율
1980년	7,969,201	1,209,806	15.2	382,743	4.8
2000년	14,311,807	134,364	0.9	2,224,433	15.5
2018년	20,499,543	52,738	0.3	5,848,594	28.5

최근 증가추세인 1인 가구 비율은 1980년 4.8%, 2000년 15.5%, 2018년 28.5%로 1980년과 비교하면 23.7%p 증가했다.

정부의 지속적인 노력에도 불구하고 출산율은 감소하고 있고 생산연령인구(15~64세) 감소 및 고령인구(65세 이상) 증가 등 인구구조가 가까운 미래에 더 큰 변화로 다가올 것을

예측하고 있다. 통계가 보여주고 있는 미래를 주의 깊게 살펴보고 대비해야 할 것이다.

[지표출처: 인구대비 혼인건수, 합계출산율, KOSIS>인구·가구>인구총조사>
가구부문>총조사가구>시도 세대구성 가구원수별 가구, KOSIS>인구·가구>
인구총조사>가구부문>총조사가구>시도/가구의 크기/핵수별 가구, 1인가구 비율]

▶ 어느 지역이 학급당 학생 수가 가장 적을까?



사회생활은 사람이 사회의 일원으로서 모여서 질서를 유지하며 살아가는 공동생활을 말한다.

아이가 태어난 뒤 첫 사회생활은 대부분 유치원 혹은 초등학교에서 시작하게 된다.

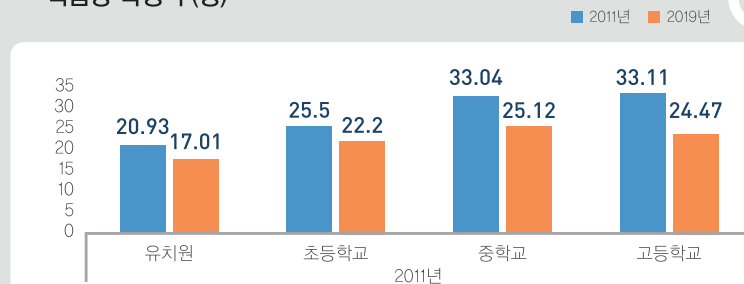
학급당 학생 수는 교육여건의 수준을 판단할 수 있는 대표적인 지표로 학급당 학생 수가 적을수록 교사와 학생 간의 긴밀한 상호작용이 가능하고 이를 통해 교육이 내실 있게 이루어질 가능성이 높아진다고 볼 수 있다.

2019년 시군구별 학급당 학생 수는 신안군이 9.15명으로 가장 적었고, 김제시 14.78명, 부산 영도구 19.72명으로 나타났다. 2011년과 비교해보면 전체적으로 학급당 학생 수는 감소했으며, 특히 고등학교의 학급당 학생 수가 가장 큰 폭으로 8.64명 감소하였다.

가장 크게 감소한 지역은 경기도 과천시로 9.06명이 감소했으며, 유치원 경기도 광명시(△8.68명), 초등학교 대전광역시 대덕구(△6.67명), 중학교 인천광역시 동구(△14.89명), 고등학교 경기도 안산시(△14.21명)로 나타났다.

학급당 학생 수가 감소하여 교육여건이 개선되는 것으로 볼 수 있으나 학생이 늘어나 교실이 부족하고 선생님이 부족한 대한민국이 될 날을 꿈꾸고 싶다.

학급당 학생 수(명)



행정 구역별	2011년					2019년				
	합계	유치원	초등학교	중학교	고등학교	합계	유치원	초등학교	중학교	고등학교
전국	28.35	20.93	25.50	33.04	33.11	22.53	17.01	22.20	25.12	24.47
경기도 과천시	32.56	18.85	31.31	37.22	33.67	23.50	11.29	25.46	28.64	21.84
경기도 광명시	31.78	24.36	29.86	35.61	35.46	24.24	15.68	25.03	28.20	24.27
대전광역시 대덕구	27.30	19.30	25.29	32.76	32.50	19.83	15.42	18.62	23.36	23.38
인천광역시 동구	27.31	20.71	24.45	35.82	29.62	21.02	16.95	20.94	20.93	23.93
경기도 안산시	32.17	20.38	29.58	36.53	38.30	23.85	15.59	24.63	27.82	24.09

[지표출처 : e-지방지표>교육>학급당 학생수(유치원·초·중·고등학교)]

▶ 고용률은 서귀포시가 최고!



대부분의 사람들은 생계를 꾸려나가기 위한 수단으로서 일정한 직업을 갖기를 원한다. 얼마나 많은 사람들이 실제로 취업하고 있는지를 가장 간명하게 나타내는 지표로 고용률이 사용되고 있으며, 고용률은 만 15세 이상 인구 중 취업자가 차지하는 비율을 나타낸다.

고용률은 노동시장의 현황을 핵심적으로 보여주는 지표라고 할 수 있다. 예전에는 고용 관련 지표로 경제활동참가율을 주로 사용했으나 실업자와 비경제활동인구의 구분이 분명하지 않아 최근에는 고용률을 많이 사용하고 있다.

시단위 중 고용률이 가장 높은 지역은 2013년 조사가 시작된 이래 서귀포시가 가장 높았으며, 성별로는 남자는 거제시가 2016년 상반기까지 높다가 조선업 불황 등으로 2016년 하반기부터는 당진시가 높게 나타났다. 여자는 조사 시작 이후 서귀포시가 가장 높게 나타났다.

군단위에서는 2013년 상반기 장수군이 남녀 모두 높게 나타났으나, 2013년 하반기부터는 남녀 모두 울릉군에서 가장 높게 나타났다.

고용률이 노동시장의 현황을 보여주는 지표라면 상용직 비중 지표는 고용시장의 안정성을 파악하기 위한 지표라고 할 수 있다. 전체 임금근로자 중 상용직으로 일하는 근로자의



비율로 상용근로자 중 고용계약설정자는 고용계약기간이 1년 이상인 경우, 고용계약미설정자는 소정의 채용절차에 의해 입사하여 인사관리 규정을 적용받는 사람으로 정의한다.

상용직 비중은 시도단위에서는 거제시가 높은 비중을 차지하다 2017년 상반기부터 화성시와 과천시와 높은 비중을 차지했으며, 군단위에서는 음성군, 칠곡군, 진천군, 울진군에서 상용직 비중이 높은 것으로 나타났다. 하루 빨리 노동시장 및 고용시장의 안정화가 이뤄지길 바라본다.

* 상용직 비중

행정 구역별	2013 1/2	2013 2/2	2014 1/2	2014 2/2	2015 1/2	2015 2/2	2016 1/2	2016 2/2	2017 1/2	2017 2/2	2018 1/2	2018 2/2	2019 1/2
화성시										86.9		86.1	85.2
과천시						85.8			83.7		87.9		
거제시	85.2	86.5	87.0	91.4	88.4		85.8	84.8					
칠곡군												81.1	81.2
음성군	82.6	81.5	78.7	80.7	80.1	82.6			82.5				
진천군										83.9	82.0		
울진군							80.3	80.0					

[지표출처 : KOSIS>고용·임금>고용>지역별 고용조사(고용률),
e-지방지표>고용과 노동>상용직 비중]



▶ 65세 이상 인구비율은 어느 지역이 가장 낮을까?

인간의 수명이 연장되면서 100세 시대가 도래했다. 유엔보고서는 평균수명이 80세를 넘는 국가가 2000년에는 6개국에 불과했지만, 2020년에는 31개국으로 급증할 것으로 예상하며 이를 '호모 헨드레드 시대'로 정의했다.

이와 함께 단순히 오래사는 리빙롱거(living longer)가 아닌 건강하게 잘 사는 리빙웰(living well)이 대두되고 있다.

EQ-5D지표는 건강관련 삶의 질 5가지 차원(운동능력, 자기관리, 일상활동, 통증/불편, 불안/우울)의 기술체계를 종합한 지표로 수치가 1에 가까울수록 삶의 질이 높다고 측정하는 지표이다.

시단위에서는 2008년 김포시가 0.982로 가장 높았으며, 2018년에는 사천시가 0.976으로 가장 높게 나타났다. 군단위에서는 2008년 강화군이 0.973으로 가장 높았으며, 2018년에는 창녕군이 0.980으로 가장 높게 나타났다. 구단위에서는 2008년 서울 마포구가 0.991로 가장 높았으며, 2018년에는 서울 동대문구가 0.979로 가장 높게 나타났다.

EQ-5D지표와 함께 주관적 건강수준인지율은 개인의 건강에 대한 주관적 인식을 파





악한 지표로 주관적 건강수준을 ‘매우 좋음’ 또는 ‘ 좋음’으로 응답한 사람의 비율로 측정한다.

시단위에서는 2008년 김제시 76.0으로 가장 높았고 2018년에는 구리시 57.5로 가장 높게 나타났으며, 군단위에서는 2008년 서천군 65.6으로 가장 높았고 2018년에는 구례군 62.3으로 가장 높게 나타났고, 구단위에서는 2008년 서울 마포구 69.8로 가장 높았고 2018년에는 대구 중구 66.3으로 가장 높게 나타났다. 10년 전과 비교했을 때 시군구 지역에서 주관적 건강수준 만족도가 전반적으로 낮아졌음을 볼 수 있다.

100세 시대는 기존의 장수, 생명연장과는 다른 의미의 과제들을 고민하게 했다. 나이 들어가는 것은 거스를 수 없지만 건강하고 아름답게 나이 들고, 품위 있는 죽음을 맞이하고 싶다는 욕구가 생긴 것이다.

관심이 높아진 만큼 질 높은 삶(well being), 잘 늙는 삶(well aging), 품위 있는 죽음(well dying)을 키워드로 하는 강의, 도서 등이 넘쳐나고 있다.

우리나라 고령인구비율은 2008년 10.23%에서 2018년 14.76%로 약 1.5배 증가했으며 2000년 7.03%와 비교하면 2배 넘게 증가했다.

고령인구비율은 전체인구수 중 65세 이상 인구의 비율로 나타내는데 울산광역시 동구가 2000년부터 2009년까지 가장 낮았고(2.74%~5.06%), 2010년부터 2018년까지는 울산광역시 북구가 가장 낮은 것(5.25%~7.26%)으로 나타났다.

고령인구비율이 가장 높은 지역과 가장 낮은 지역의 차이는 2000년 18.18%p에서 2008년 25.87%p, 2018년 31.61%p로 점점 격차가 커지고 있다.

65세 이상 고령인구 중 1인 가구의 비율인 독거노인가구비율은 고령화사회에서 가장 중요한 노인 생활의 질을 보여주는 지표로 혼자 사는 노인의 경우 대부분 사회적으로 고립되어 경제적인 측면뿐만 아니라 사회관계 측면에서도 어려움에 직면하고 있어 노인들의 삶의 질 측정에 중요한 지표로 평가되고 있다.

우리나라 독거노인가구 비율은 2000년 3.8%에서 2018년 7.2%로 89.5% 증가했다. 지역별로 살펴보면 2000년 인천 연수구가 1.1%로 가장 낮았으며 2018년 대전광역시 유성구가 3.3%로 가장 낮았다. 급속도로 증가하고 있는 고령인구 및 독거노인가구비율에 따른 대비책 등은 지속적으로 고민하고 연구해야 할 과제임에 틀림없을 것이다.

[지표출처 : e-지방지표>건강>EQ-5D, e-지방지표>인구>고령인구비율,
e-지방지표>가족>독거노인가구비율]



나도 R유저 ⑥



R에 도전하자... 따라가다보면, 나도 R유저⑥



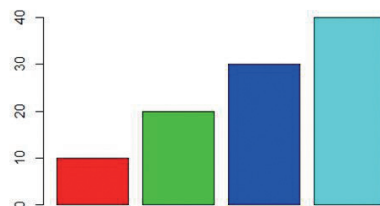
‘R에 도전하자 ①’부터 ⑤에서 R의 설치부터 시작하여 몇 가지 기능을 따라 해보고 R에서의 색깔처리, R의 기초적인 통계함수 및 R의 벡터 및 데이터 프레임을 사용한 자료처리, 함수작성 등에 대해 알아보았다. 이번 호에서 R 그래픽 장치 및 여러 개의 그래프를 다루는 방법에 대해서 알아보기로 하자.

1 | R의 그래픽 장치(graphics device)

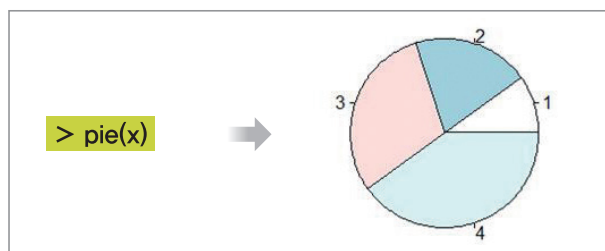


이전 호의 R에 도전하자에서 plot, pie 명령들을 사용한 보기가 있는데 이 경우 별도의 그래픽 장치를 선택하지 않고 그래픽을 구현하였다. 예를 들어

```
> x <- c(10,20,30,40)
> barplot(x, col=2:5)
```



처럼 `barplot` 명령을 하면 바로 그래픽 창이 열리면서 그래픽이 만들어진다. 이 경우 다른 그래픽 명령을 사용하면 위의 막대그래프가 사라지고 그 창에 새 그래픽이 만들어진다. 예를 들어 같은 자료로 원그래프를 그리면



로 앞의 막대그래프 자리에 원그래프가 그려져 두 그래프를 비교하려고 한 경우 두 그래프가 동시에 그려지지 않아 어려움에 봉착한다.

R은 여러 가지 그래픽 장치를 제공하는데 그래픽을 사용할 때 기본값으로 그래픽 창이 열리는 것으로 지정되어 있어서 위의 예처럼 그래픽 창이 없으면 새로 열고, 기존 창이 있으면 기존 창에 그리게 된다 (뿔어 씹). R이 제공하는 그래픽 장치는 크게 세 가지로 분류할 수 있다.

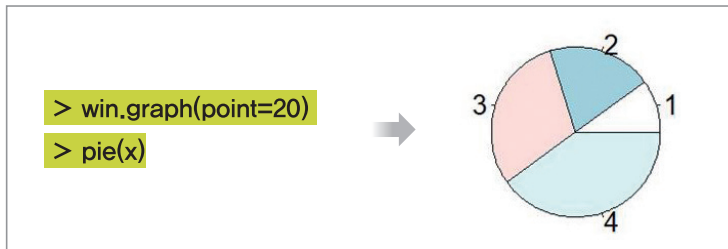
1. `win.graph` 는 R에서 바로 창을 띄워 그래픽을 구현하게 되고,
2. `bmp`, `jpeg`, `png`, `tiff`는 각각 BMP 형식, JPG 형식, PNG 형식 및 TIFF 형식의 파일로 그래픽을 저장하게 된다.
3. `PDF`, `PostScript` 형식이 있는데 이 두 형식은 앞의 형식과 달리 생성된 그래픽이 여러 페이지에 저장될 수 있다. 즉, `win.graph`, `bmp`, `jpeg`, `png`, `tiff`는 한 장치에 한 개의 그래픽만 들어가나 `pdf`나 `postscript` 장치는 두 개 이상의 그래픽을 만들면 각각의 그래픽이 한 페이지가 되어 여러 개의 그래픽을 여러 페이지에 저장할 수 있다.

win.graph 함수와 윈도우 그래픽 장치

`win.graph` 함수는 그래픽 장치로 R에서 새 그래픽 창을 생성하는 함수이다. 이 함수는

```
win.graph(width, height, pointsize)
```

로 사용하며 `width`와 `height`는 열릴 그래픽 창의 크기를 인치 단위로 설정하고 기본값은 둘 다 7이다. `win.graph`의 결과가 대화형 창이 열리는 것이라서 마우스 드래그를 통해 창의 크기를 조정할 수 있으므로 `width`와 `height`는 대개 별도로 설정하지 않는다. `pointsize`는 글자 크기인데 기본값은 12(포인트)이다. 예를 들어 위의 원그래프를 그리기 위해



와 같이 win.graph에 pointsize에 20을 설정하면 글자 1, 2, 3, 4의 크기가 20 포인트로 커졌음을 앞선 원그래프와 비교해 보면 알 수 있다. 두 개 이상의 창을 열려면 win.graph 명령을 두 번 이상 하면 된다. 즉,

```
> win.graph()
```

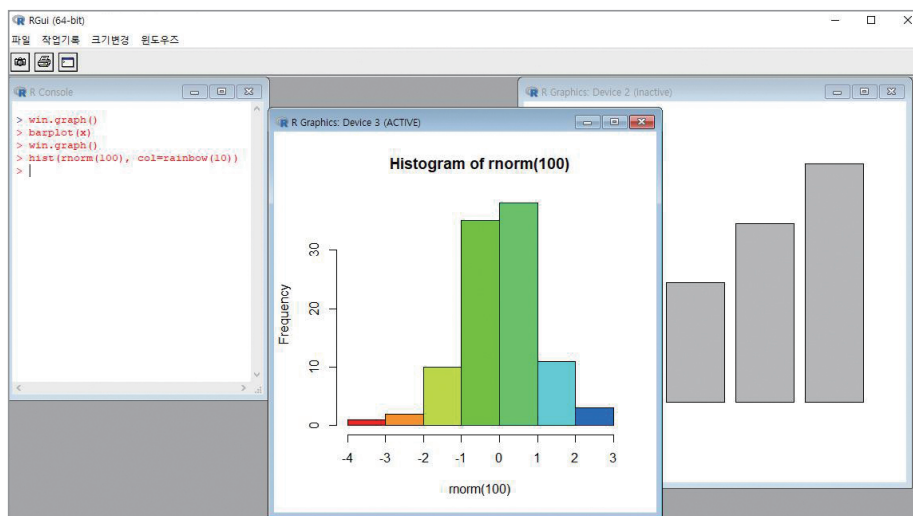
```
> barplot(x)
```

```
> win.graph()
```

```
> hist(rnorm(100), col=rainbow(10))
```



명령의 결과는 처음의 win.graph에 의해서 창이 열리고, 이 창에 barplot을 그리고, 두 번째 win.graph 명령에 의해 새 그래픽 창이 하나 더 열리게 되며 이 창에 hist 명령에 의해 정규분포의 난수 100 개에 대한 히스토그램을 그리게 된다. 결과는 다음과 같이 두 개의 그래픽 창이 열리고 두 창에 각각의 그래픽이 생성되었다.





그래픽 장치 닫기 - dev.off와 graphics.off

생성된 그래픽 장치를 닫는 것은 dev.off나 graphics.off 함수로 실행한다. 이 함수에 매개변수를 사용하지 않으면 가장 최근 그래픽 장치를 닫으며, 매개변수로 그래픽 장치 번호를 주면 해당 번호의 그래픽 장치를 닫는다. 위의 그림에서 그래픽 창의 타이틀 바(title bar)를 보면 그래픽 장치의 번호가 2와 3이 사용되었음을 알 수 있는데, 이 상태에서

```
> dev.off( )
```

또는

```
> dev.off(3)
```

를 명령하면 3번 장치의 히스토그램 창이 닫히고,

```
> dev.off(2)
```

를 하면 2번 장치의 막대그래프 창이 닫힌다. win.graph의 경우 대화형 창이 생성되는 장치이므로 마우스로 해당 창을 닫는 것이 더 편리할 수 있어서 dev.off 함수로 개별 창을 닫지 않아도 되나 뒤에서 소개할 다른 장치들은 이 함수를 사용하는 것이 좋다.

열려진 모든 그래픽 장치를 한 번에 닫는 명령으로 graphics.off 함수를 사용할 수도 있다. 위의 창에서

```
> graphics.off( )
```

를 명령하면 두 창이 한꺼번에 닫힌다. dev.off와 graphics.off 함수는 뒤에서 소개할 장치에도 같은 방법으로 적용된다.

bmp, jpeg, png, tiff 함수

bmp, jpeg, png, tiff 함수는 이름 그대로 R 그래픽을 BMP 형식, JPG 형식, PNG 형식, TIFF 형식으로 저장한다. 이 그래픽 장치가 열린 상태에서 그래픽 명령을 실행하면 결과가 바로 지정된 파일로 저장되어 R 콘솔에는 아무 메시지도 출력되지 않으며(메세지가 출력되면 에러 메시지 등 뭔가 잘못되었을 가능성이 크다) 당연히 그래픽 창도 열리지 않는다. 요즘처럼 컴퓨터의 속도가 빨라진 경우 마치 R이 아무 것도 하지 않은 것처럼 보이기도 한다. 이들 함수는 각각

```
bmp(filename = "Rplot%03d.bmp", width = 480, height = 480, units = "px",  
      pointsize = 12, bg = "white", ...)
```

```
jpeg(filename = "Rplot%03d.jpg", width = 480, height = 480, units = "px",  
      pointsize = 12, quality = 75, bg = "white", ...)
```



```
png(filename = "Rplot%03d.png", width = 480, height = 480, units = "px",
     pointsize = 12, bg = "white", ...)
```

```
tiff(filename = "Rplot%03d.tif", width = 480, height = 480, units = "px",
     pointsize = 12, compression = c("none", "rle", "lzw", "jpeg", "zip", "lzw+p",
     "zip+p"), bg = "white", ...)
```

으로 사용법이 거의 같으며 저장 형식에 따라 한두 개의 매개변수가 더 있는 경우가 있다. 각 매개변수는 다음과 같이 설정한다.

- **filename**은 그래픽을 저장할 파일의 이름을 설정하며, 파일 이름만 준 경우 R의 현재 작업디렉토리에 파일을 만들게 되므로 특정한 경로에 가져다 두고 싶으면 파일의 전체 경로를 주거나 `setwd` 명령으로 R의 기본 작업 디렉토리를 변경하여야 한다. **filename**은 생략하면 `Rplot001`, `Rplot002` 등으로 R이 자동으로 번호를 붙여 기본 작업 디렉토리에 파일을 생성하며 파일확장자는 저장형식에 따라 `bmp`, `jpg`, `png`, `tif`로 자동생성된다.

- 파일의 경로구분: 윈도우의 경우 경로구분은 `W`(자판에서 `W`)를 사용하나 R에서 `W`는 특수문자이므로 경로구분에는 `/` 또는 `WW`(두 개 붙여 씀)를 사용한다. 예를 들어 파일의 전체 경로가

`C:\Wmydata\Wmyfolder\Wpic.jpg` 이면 이면 R에서는 `C:WWmydataWWmyfolderWWpic.jpg` 또는 `C:/mydata/myfolder/pic.jpg`를 사용한다.

- R에서 현재 작업 디렉토리는 `getwd()`를 사용하여 확인할 수 있고, 작업디렉토리를 바꾸려면 `setwd` 함수를 사용할 수 있다. `setwd` 함수를 사용하면 이후 생성되는 파일은 특별히 경로를 지정하지 않으면 `setwd`에서 지정한 경로를 사용한다.

```
> getwd()
```

```
[1] "C:/Users/Sim/Documents"
```

는 현재의 작업 디렉토리이며 이를 `D:WHWP`로 바꾸려면

```
> setwd("D:/HWP/")
```

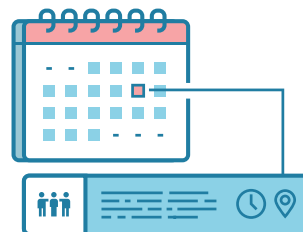
를 사용하며 결과로

```
> getwd()
```

```
[1] "D:/HWP"
```

작업 디렉토리가 변경되었음을 알 수 있다.

- `bmp`, `jpeg`, `png`, `tiff`로 그래픽 장치를 열고 둘 이상의 그림을 그리면 `Rplot001`, `Rplot002`와 같이 번호를 하나씩 증가시켜 그림을 만든다. 여기서 주의해야 할 점은 `dev.off`나 `graphics.off`로 그래픽 장치가 닫히면 이 번호도 초기화 된다.



– 위의 그래픽 장치로 파일이 만들어질 때 기존에 같은 이름의 파일이 존재하더라도 사용자 확인없이 덮어쓰므로 파일이름을 줄 때는 기존에 같은 이름의 파일을 지워도 되는지 확인하여야 한다. 파일이름을 주지 않더라도 Rplot001 등이 말없이 덮어써짐에 함께 유의하여야 한다.

- **width, height**는 만들어질 파일의 폭과, 높이를 설정하며
- **units**는 단위를 설정한다. 기본값은 픽셀이다.
- **pointsize**와 **bg**: 각각 글자크기 및 배경색을 설정한다.
- **quality**: jpeg 함수에는 quality를 추가 설정할 수 있는데 JPEG이 압축파일이라 압축의 정도를 얼마나 할 것인지 설정하는 것으로 quality가 높을수록 압축이 덜되어 파일크기는 커지고 그림의 정밀도는 낮아진다.
- **compression**: tiff 함수의 compression은 tiff 파일을 압축할 것인지, 압축한다면 어떤 방법으로 할 것인지 설정하며, 기본값은 압축하지 않는 것이다. 따라서 tiff 파일은 일반적으로 파일크기가 크다.

위의 함수를 이용하여 그래픽을 저장하는 예를 살펴보자. 이를 위해 먼저 파일을 저장할 폴더를 지정하자.

```
> setwd("D:/HWP/lecture/기고문/통계진흥원/images")
```

이제 jpeg 장치를 열어 100개의 난수를 표준정규분포에서 만들어 이를 hist.jpg에 저장하려면 아래와 같이

```
> jpeg("hist.jpg")
```

```
> hist(rnorm(100), col=rainbow(10))
```

명령해보자. 파일 탐색기를 통해 이 파일이 생성되었음을 확인할 수 있으나 이 파일은 R 그래픽에 의해 열려있는 상태라 다른 프로그램은 아직 이 파일을 열지 못한다. 따라서

```
> dev.off( )
```

로 장치를 닫아주면 결과가 저장되고 다른 그래픽 앱(GIMP, irfanview, 포토샵 등)으로 이 파일을 열거나 편집 등을 할 수 있다. 이번엔 파일 이름 없이 PNG 장치를 열어보자. 이를 위해

```
> png( )
```

로 장치를 열고 두 개의 그래픽 명령을 사용하면

```
> hist(rnorm(100), col=rainbow(10), freq=F)
```

```
> curve(dnorm, from=-3, to=3)
```

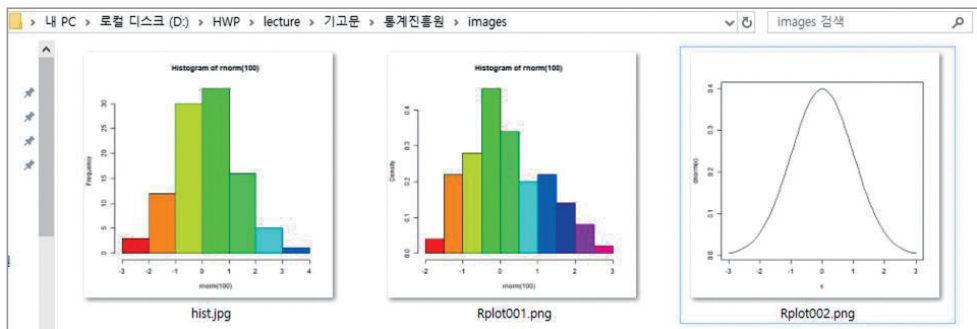
파일 탐색기에서 이 파일이 생성된 것을 확인할 수 있으며, 더 이상 다른 그림을 그리지 않으며 다음 명령으로 장치를 닫으면 된다.

```
> dev.off( )
```

이 경우 PNG 장치를 파일 이름 없이 열었으므로 파일명은 Rplot001.png 등이 되며, 두 개의 그래픽 명



령이 실행되었으므로 정확하게는 Rplot001.png와 Rplot002.png가 생성되었음을 알 수 있다. 위 명령 결과 파일탐색기를 확인해보면 다음과 같이 파일들이 생성되어 있음을 알 수 있다(파일 탐색기가 '아주 큰 아이콘으로 보기'로 설정된 결과임)



위에서 만든 세 개의 파일은 일반적인 jpg, png 파일이므로 GIMP나 포토샵 등 이미지 편집 프로그램을 사용해 편집이 가능하다. bmp나, tiff 도 저장형식의 차이를 제외하면 사용법 및 결과물을 얻는 방법은 거의 같으므로 위의 함수 jpeg이나 png와 같은 방법으로 시도해 보기 바란다.

pdf와 postscript 함수

PostScript는 한 동안 출판에서 표준으로 사용하던 형식 중의 하나로 텍스트 파일로 저장되며 PostScript language로 작성된다. 텍스트 파일이라 파일의 크기가 큰 것이 단점 중의 하나이다. 요즘은 PDF가 표준으로 많이 사용된다. 그래픽을 PDF와 PostScript 형식으로 저장하려면 각각 pdf 및 postscript 함수를 사용하며 이 두 명령은

```
pdf(file = if(onefile) "Rplots.pdf" else "Rplot%03d.pdf", width, height,
    onefile, paper, bg, fg, ... )
```

및

```
postscript(file = if(onefile) "Rplots.ps" else "Rplot%03d.ps", onefile, width, height,
    bg, fg, horizontal, paper, ...)
```

로 사용한다. 매개변수는

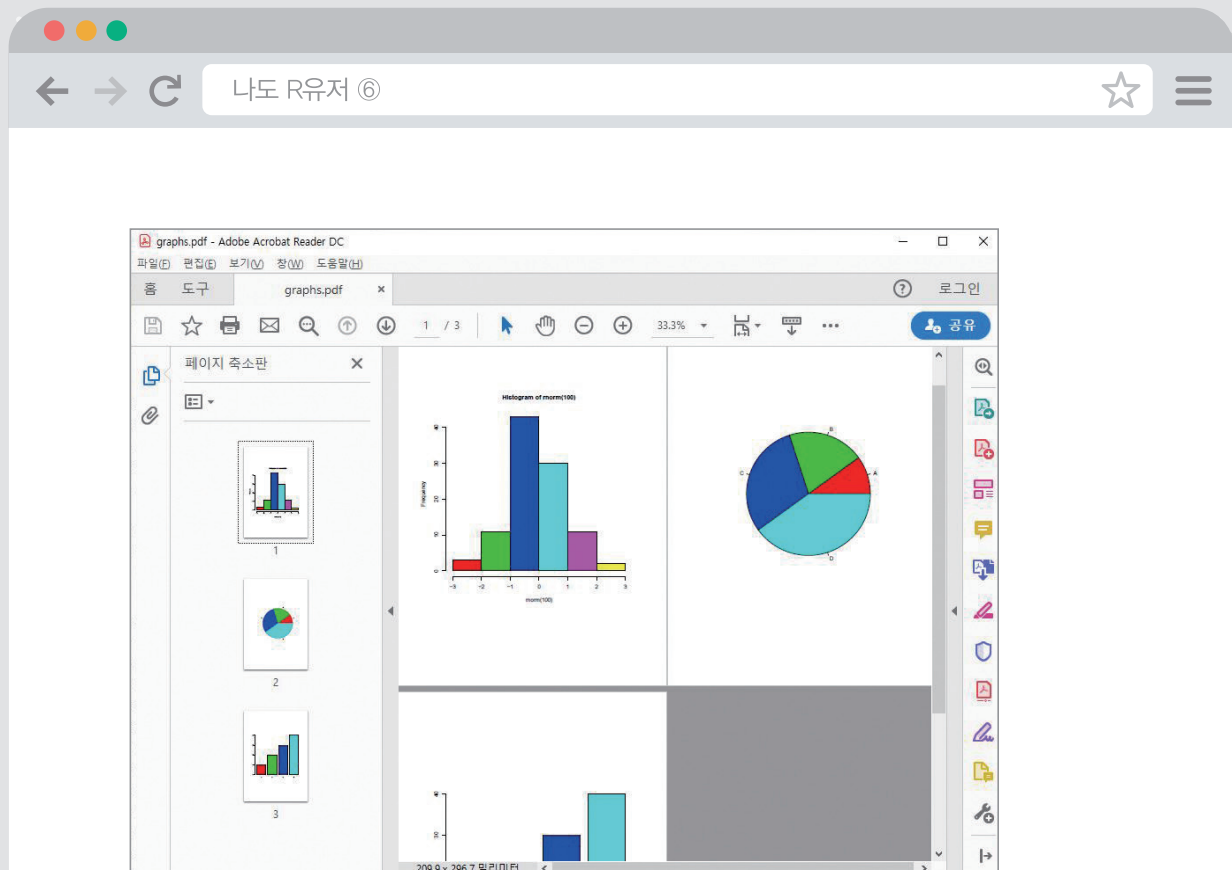
- **file:** 앞에서 마찬가지로 그래픽을 저장할 파일 이름을 설정한다. 아래의 onefile이 FALSE이면 앞의 bmp, jpeg 등의 장치와 마찬가지로 각 그래픽 명령이 있을 때마다 Rplot001, Rplot002 등으로 이름을 붙이며 각각의 파일에 한 개의 그래픽이 만들어진다. 이 값이 TRUE이면 그래픽 장치가 닫힐 때까지 한 파일에 모든 그래픽을 그리며 각 그래픽은 이 파일의 한 페이지씩 구성된다.
- **onefile:** 장치가 열린 후 둘 이상의 그래픽이 있을 때 이를 하나의 파일로 할지 각각의 파일로 할지 설정하며 기본값은 TRUE로 한 파일에 저장한다.
- **bg, fg:** 배경색과 전경색을 설정한다. 배경색은 "transparent"가 기본값, 전경색은 "black"이 기본값이다.
- **width, height:** 그림의 크기를 설정한다. 단위는 인치이다.
- **horizontal:** 그림을 가로로 할지 세로로 할지 설정한다. 기본값은 가로방향이다.
- **paper:** 종이 크기 설정이다. 가능한 값은 "a4"(A4), "letter"(미국 letter size), "legal", "executive" 또는 "a4r"(A4 가로방향) 등이다. 기본값은 "special"로 width, height 값에 주어진 크기로 만든다. "default"로 설정하면 option에 설정된 papersize 값을 사용한다. option에 사용된 papersize는 options()\$papersize로 확인할 수 있으며 운영체계에 따라 한글 윈도우의 경우 이 값은 대개 "a4"이다.

이제 pdf 함수를 사용하여 그래프를 pdf 파일에 저장하기 위해 세 개의 그래픽-히스토그램, 원그래프, 막대그래프를 다음의 명령으로 생성하는 예를 생각해 보자.

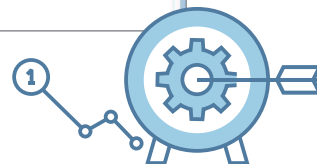
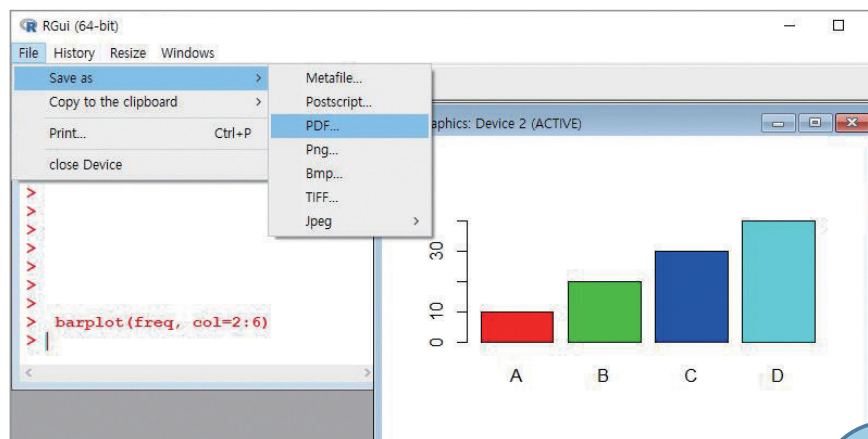
```
> freq <- c(10,20,30,40)
> names(freq) <- LETTERS[1:4]
> pdf(file="E:/HWP/lecture/기고문/통계진흥원/images/graphs.pdf", paper="a4")
> hist(rnorm(100), col=2:8)
> pie(freq, col=2:6)
> barplot(freq, col=2:6)
> dev.off()
```

위 명령의 결과로 graphs.pdf 파일이 경로 E:\HWP\lecture\W기고문\통계진흥원\Wimages에 생성되었으며, 이 파일엔 위의 hist, pie, barplot 함수의 결과가 한 페이지씩 구성되어 세 페이지짜리 PDF 파일이 생성된다. 이 파일을 Adobe PDF Reader로 열어보면 다음과 같다.





참고 R에서 대화형으로 그래프를 생성한 후(즉, win.graph 장치에 그래프를 만든 후) 그래픽 창을 활성화하고 File → Save as를 선택하면 다음의 그림에서 보는 것처럼 윈도우에 보이는 그래프를 metafile(클립보드), PostScript, PDF, png, jpg, bmp, TIFF 등으로 저장할 수 있다.



또한 그림영역에서 마우스 오른쪽 버튼을 클릭하여도 클립보드로 복사하거나 일부 형식으로 그림을 저장할 수 있다. 그렇다면 위의 함수들은 굳이 사용할 필요가 있을까 의문을 가질 수 있다. 하지만 대화형으로 하나씩 분석하는 경우에는 문제가 없으나 함수 작성 등을 통해 한 자료에 대해서 한꺼번에 여러 명령을 수행하는 경우 그래프를 별도로 저장하여 분석이나 보고서 등에서 사용하는 것이 편하므로 위의 그래픽 장치 함수를 이해하고 구현할 수 있어야 한다.

2 | 한 장치에 두 개 이상의 그래프 그리기 - par 함수와 mfrow, mfcoll



현재까지 그래픽은 한 화면에 하나씩 그리기만 알아보았는데 R은 한 그래픽 화면을 표의 칸처럼 나눠 각 칸에 그래픽을 하나씩 그리는 방법을 제공한다. 예를 들어 그래픽 영역을 상하 두 개로 나누어 위아래에 각각 다른 그림을 그리거나, 화면을 2개의 행, 2개열을 가진 4개의 영역으로 나누어 4개의 그림을 그리는 것이 가능하다. 이 방법은 그래픽 매개변수를 변경하는 간단한 방법으로 가능하다. 그래픽 변수의 변경은 par 함수를 사용하고, 이 함수의 매개변수 중 mfrow 또는 mfcoll 값을 바꾸어 줌으로 가능하다. mfrow, mfcoll은 그래픽 영역이 분할할 때 사용할 행의 개수와 열의 개수 두 값을 인수로 갖는다. 기본값은 c(1,1)이며 설정하는 방법은 c(nrow, ncol) 형식으로 설정한다. 기본값이 c(1,1)이므로 행도 1개 열도 1개이기 때문에 그래픽 영역에 그림을 하나만 그리게 된다.

c(nrow, ncol)에서 nrow와 ncol은 그래픽 영역을 칸으로 나눌 때 각각 행의 개수와 열의 개수를 의미한다. mfcoll은 그림의 순서가 열우선이며, mfrow는 그림순서를 행우선으로 한다. 다음은 각각의 mfcoll 또는 mfrow의 설정에서 그림 영역이 나누어진 형태와 그림의 순서를 그림으로 표시한 것이다.

mfcoll=c(1,2)	mfrow=c(2,1)	mfcoll=c(2,2)	mfrow=c(2,2)												
<table><tr><td>1</td><td>2</td></tr></table>	1	2	<table><tr><td>1</td></tr><tr><td>2</td></tr></table>	1	2	<table><tr><td>1</td><td>3</td></tr><tr><td>2</td><td>4</td></tr></table>	1	3	2	4	<table><tr><td>1</td><td>2</td></tr><tr><td>3</td><td>4</td></tr></table>	1	2	3	4
1	2														
1															
2															
1	3														
2	4														
1	2														
3	4														

par 함수로 그래픽 설정을 변경한 경우, 다시 변경할 때까지 변경된 설정이 계속 유지되므로 par 함수로 설정을 변경할 때는 그림을 그린 후 그래픽 설정을 이전의 값으로 되돌려 두는 것이 권장된다. 이를 위해서는 대개 원래의 값을 임시로 저장했다가 그래픽 명령을 실행한 후 임시로 저장했던 값을 다시 par 함수를 사용하여 설정하는 방법을 사용한다. 예를 들어 mfrow를 변경한 경우



```
omfrow <- par("mfrow")
```

그래픽 명령들

```
par("mfrow"=omfrow)
```

로 현재의 mfrow 값을 omfrow(old mfrow라는 의미)에 저장한 후 그래픽 명령들을 실행한 뒤 다시 par 함수로 mfrow의 값을 원래 값으로 되돌린다. 함수 내에서 이 방법을 사용할 때는 대개 on.exit 함수를 사용하여 함수가 끝날 때 자동으로 동작되도록 한다.

이제 한 그래픽 영역에 둘 이상의 그림을 넣은 예를 생각해보자. 다음 명령

```
> omfrow <- par("mfrow")
```

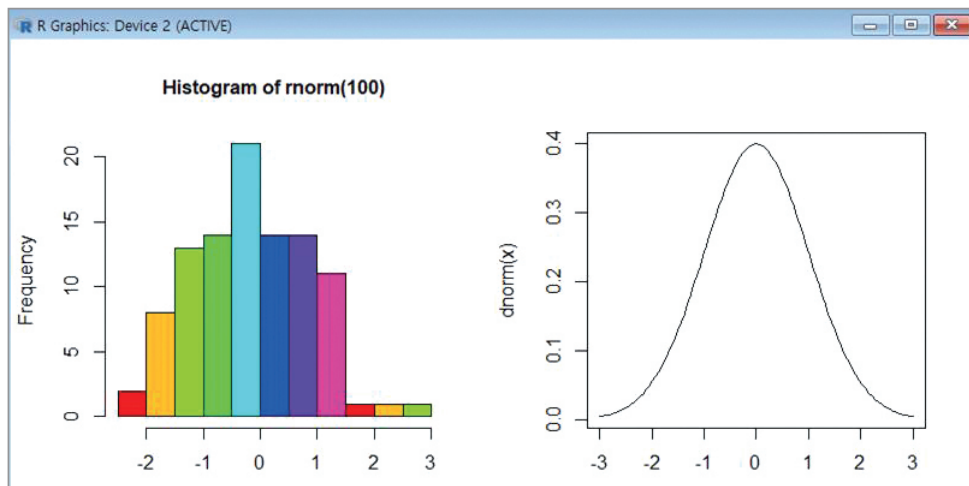
```
> par(mfrow=c(1,2))
```

```
> hist(rnorm(100), col=rainbow(8))
```

```
> curve(dnorm, from=-3, to=3)
```

```
> par("mfrow"=omfrow)
```

의 결과는 아래의 그림과 같이 한 그림영역에 두 개의 그림을 가로로 그리게 된다. 위 명령의 첫 줄과 마지막 줄은 각각 현재의 그래픽 변수값을 저장하고 그림 그리기가 끝난 후 원래의 설정으로 바꾸어 주는 과정이며 mfrow가 c(1,2)로 설정되었으므로 하나의 행, 두 개의 열로 한 화면이 구성되어 그림영역이 좌우로 반반씩 나누어짐을 알 수 있다.



이번엔 그래픽 영역을 2로 나누되 한 번은 mfrow를, 한번은 mfcpl을 설정해보자. 이렇게 4개의 그래픽 영역을 만들고 평균인 10, 20, 30이고 분산은 모두 1인 난수를 1,000개씩 생성해서 히스토그램을 그려



보자. 그래픽 영역은 4개이나 일부분에만 그릴 수 있음을 확인하기 위해 그림은 3개만 그려 보도록 한다. 먼저 mfrow를 사용하여

```
> omfrow <- par("mfrow"); omfcol = par("mfcoll")  
> par(mfrow=c(2,2))  
> for (i in 1:3) hist(rnorm(1000, i*10, 1), col=rainbow(8))
```

세 개의 히스토그램을 그린 후

```
> win.graph()
```

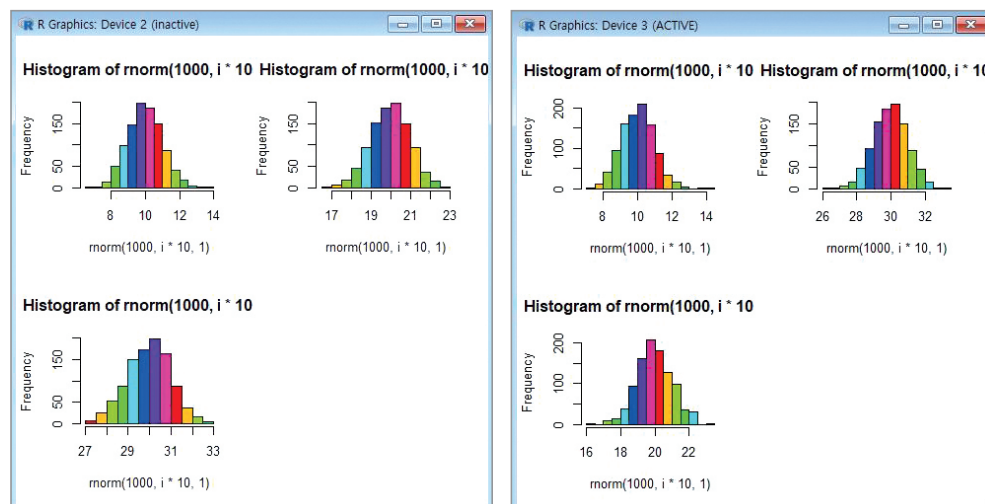
로 새 그래픽 창을 생성한 다음

```
> par(mfcol=c(2,2))  
> for (i in 1:3) hist(rnorm(1000, i*10, 1), col=rainbow(8))  
> par("mfrow"=omfrow, "mfcoll" = omfcol)
```



명령한다. 위 명령의 결과는 아래의 그림이며 왼쪽 그림은 mfrow를 설정하여 그림의 순서가 행 우선으로 채웠고, 오른쪽 그림은 열 우선으로 채웠음을 x 축의 값을 참조하여 평균값이 10, 20, 30이 되는 그림으로 알 수 있다.

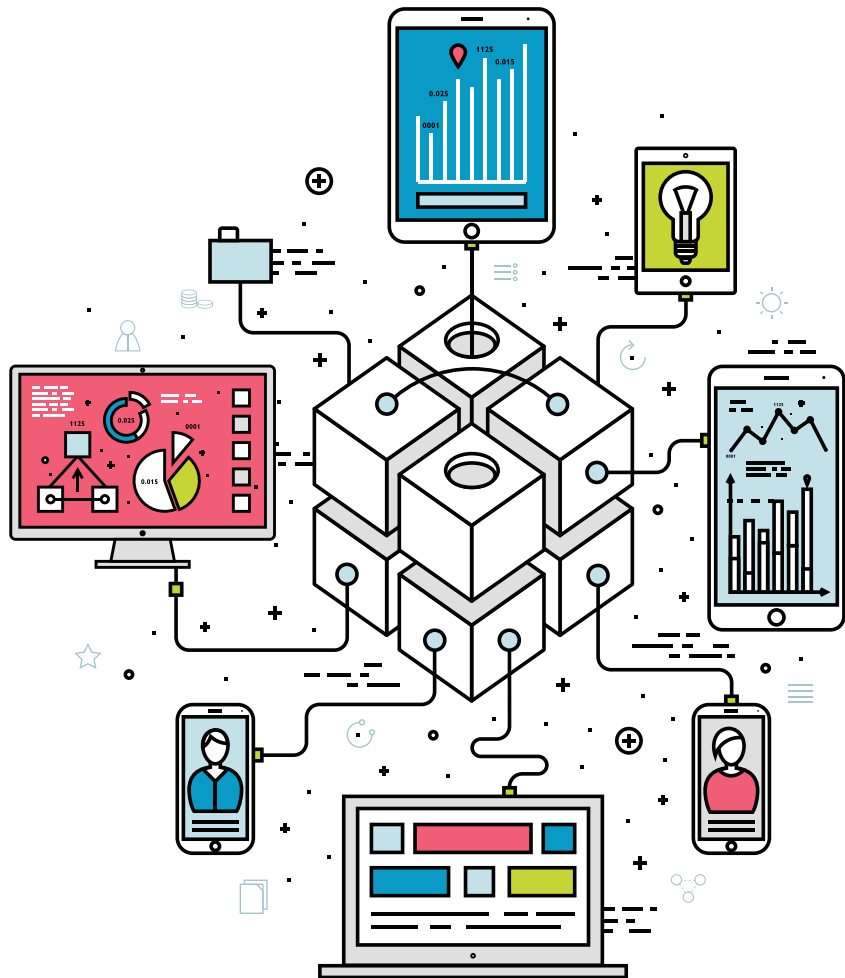
위에서 사용한 for는 반복문으로 1:3, 즉, i의 값이 1, 2, 3일 때 hist(rnorm(1000, i*10, 1), col=rainbow(8)) 각각 한 번씩 실행하게 되어 평균이 각각 i*10=10, 20, 30일 때의 정규분포 난수를 생성하여 히스토그램을 그리게 된다.



인포그래픽 나도 할 수 있다 series 11

실전에서 자주 사용하는 엑셀 기반의 서열데이터& 발표용 통계 수치 시각화 방법

공공기관을 비롯해 기업과 언론에서 가장 많이 사용하는 정보의 종류가 바로 숫자 기반의 데이터다. 빅데이터, 생활 통계 활용 교육 등의 시대 흐름에 따라 통계 인포그래픽이 많이 사용되고 있다. 통계는 데이터에서 어떤 메시지를 추출해 설명하느냐에 따라 표현 방법이 달라질 수 있다. 이번 호에서는 실무에서 자주 사용하는 엑셀 기반의 통계 자료, 즉 서열 형태의 데이터를 임의로 추출하여 시각화하는 방법과 간단한 발표용 통계 수치를 주목도를 높이기 위해 시각표로 만드는 과정을 살펴보고자 한다.



Infographics



① 서열 정보 분석법

서열은 말 그대로 데이터의 순서가 정해지는 경우다. 숫자기반 통계 데이터 해독의 핵심은 바로 먼저 데이터에서 독립변수와 종속변수값을 찾는 것이다. 독립변수는 문장에서 주어나 마찬가지로 기준이 되는 변수다. 지금부터 데이터 정보 기획에서 중요한 변수 찾기와 데이터 편집 과정을 알아보자.

▶1단계(읽기) 제시 데이터 상관관계를 파악한다.

표는 숫자의 관계를 일정 기준에 의해 정렬한 것을 말한다. 먼저 각 열에 있는 항목과 각 행에 있는 데이터의 흐름을 파악하는 것이 중요하다. 열의 개수와 행의 개수, 최댓값과 최솟값, 열 구분 기준 등은 모두 유심히 봐야 할 지표다. (그림1) 표는 광고선전비 상위 지출 기업을 나타낸 표다.

그림 1

(단위 : 백만 원, %)				
순위	상장 여부	회사명	광고선전비	증감률
1	상	삼성전자	734,815	-26.1
2	상	LG전자	528,545	-0.6
3	상	현대자동차	374,488	6.0
4	상	LG유플러스	300,222	5.2
5	상	아모레퍼시픽	264,464	-5.6
6	상	SK텔레콤	213,605	-10.0
7	상	LG생활건강	186,574	7.7
8	상	KT	162,953	0.3
9	상	롯데쇼핑	159,945	-21.5
10	비	동서식품	159,941	-10.5
11	상	에이치엘비	155,728	15.9
12	상	한국타이어	151,267	1.2
13	상	기타자동차	138,508	-13.9
14	상	KT&G	132,647	58.9
15	상	삼성물산	126,068	6.0
16	상	영리정보시스템	119,415	52.8
17	상	롯데칠성음료	106,850	-26.5
18	상	LG디스플레이	106,417	11.4
19	상	대한항공	98,935	-7.9
20	상	종교기업은행	93,619	-9.9
21	상	남양유업	90,790	-4.6
22	상	포스코	85,486	31.0
23	상	메리유업	69,036	17.1
24	상	동원F&B	67,000	1.9
25	비	해나은행	65,219	-0.9
26	상	LG	64,286	31.5
27	상	글로벌이머	62,658	17.5
28	상	오뚜기	60,602	3.1
29	상	농심	59,992	13.3
30	비	신한은행	55,996	-4.3
합계			4,997,063	

▲ 2014 회계연도 광고선전비 30위 기업 현황(출처 : 재방닷컴)

분석 P·O·I·N·T

- 독립변수, 종속변수를 결정한다.
- 데이터에서 추출할 순위를 결정한다.
- 버려야 할 항목을 결정한다.
- 그래프 유형을 결정한다.

▶2단계(요약) 표를 읽고 데이터를 재구성한다.

해당 표는 상위 30개 기업의 광고선전비, 증감률, 상장 여부 등 3가지의 종속변수를 가지고 있다. 데이터를 요약할 때 버려야 할 항목을 결정하는 것이 필요하다. 해당 자료의 포인트는 이 중 광고선전비 규모를 파악하는 것이다.

표를 텍스트로 재구성

임의로 전체 데이터에서 상위 Top 10 자료를 추출한다.

그림 2

독립변수	종속변수
기업명(10개)	광고선전비, 증감률

분석 P·O·I·N·T

- 독립변수 결정 : 상위 10개 기업명
- 종속변수 결정 : 10개 기업 광고선전비, 증감률
- 그래프 유형 : 막대그래프(10개 기업의 광고선전비 순위)

※ 이 외의 자료는 참조만 하고 채택하지 않는다.



▶3단계(레이아웃) 그래픽 제작을 위한 최종 정보 요약 및 레이아웃 결정

막대그래프 중 수평 막대그래프와 수직 막대그래프 중 하나를 결정해야 한다. 영향을 주는 것은 바로 X축을 이루는 독립변수의 개수다. 즉 막대그래프 수가 많은 경우 수평 막대그래프로 그리는 것이 좋다. 최근 모바일 기반으로 정보를 요약하는 경우가 많아 스마트폰의 제한된 가로폭을 감안하면 세로막대그래프보다 가로막대그래프가 자주 사용되는 것을 알 수 있다. 10개 기업 '광고선전비'를 순차 배열로 표시해야 하는데 이때 위에서부터 순서대로 나열하는 것이 중요하다. 10개의 수평 막대그래프를 그릴 때는 1~5위, 6~10위 사이에 쉬어가는 가로선을 그어 주면 데이터를 좀 더 쉽게 읽을 수 있다. 홍보를 위해서 제목을 다는 경우 검색엔진에서 잘 노출되도록 글쓰기를 하는 것을 잊지 말아야 한다.

제목 : 광고선전비 상위 10개 기업

그림 3

광고선전비 상위 10개 기업을 조사한 결과 삼성전자의 광고선전비가 지난해 7천348억 원으로 1년 전 9천943억 원에서 26.1% 감소했다. 2위는 LG전자 5천288억 원(전년 대비 0.6% 감소), 3위 현대자동차 3천745억 원(전년 대비 6.0% 증가), 4위 LG유플러스 3천2억 원(5.2% 증가), 5위 아모레퍼시픽 2천645억 원(5.6% 감소), 6위 SK텔레콤 2천136억 원(10.0% 감소), 7위 LG생활건강 1천866억 원(7.7% 증가), 8위 KT 1천630억 원(0.8% 증가), 9위 롯데쇼핑 1천599억 원(21.5% 감소), 10위 동서식품 1천598억 원(10.5% 감소)이다. 30개 기업의 광고선전비는 2013회계연도 5조2천211억 원에서 지난해 4조9천971억 원으로 4.3% 감소했다.

(자료 : 재방닷컴, 데이터 : 2014 회계연도 기준)

분석 P·O·I·N·T

- 수평 막대그래프의 경우 Y축이 X축이 되며, 반대로 X축은 Y축이 된다.
- 웹, 리포트 등에 사용 가능하도록 세로형 사각형으로 형태를 결정한다.
- 하단 부분에 출처, 작성기준 등을 명시한다.

▶ 4단계(제작 시 고려해야 할 사항)

수평 막대그래프를 그릴 때 숫자의 단위는 읽기 쉽도록 통일한다. 증감률에서 증가와 감소를 나타내는 컬러를 사용하는 것도 아이디어다.

💡 분석 P·O·I·N·T

- 수평 막대그래프는 언제, 어떻게 사용해야 하는지 기본 개념을 익힌다.
- 건조한 느낌의 그래프에 한두 컷의 관련성 높은 이미지를 넣어본다.(잉여 공간에 배치)
- 기업 명칭은 가급적 정확하게 사용해야 한다. 작성 후 기업명, 데이터에 오류는 없는지 다시 검토한다.

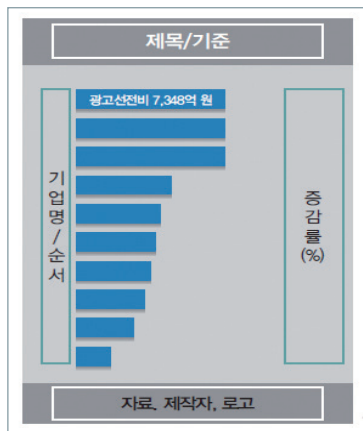


그림 4



그림 5 최종 인포그래픽 제작 사례 모습(출처: 인포그래픽기획&디자인By 파워포인트/ 이수동, 이해강, 예문사)

▶ 최종 인포그래픽 제작 사례

회사별 광고 선전비를 표현하기 위해 광고 매체 중 대표적인 TV 모양 안에 각 회사별 데이터를 넣는 방법을 사용하였다. 금년도 및 전년도 대비를 표현하기 위해 도형을 이용해 등근 막대 그래프로 표현하였다. 주목도를 높이기 위해 막대 그래프 색을 순차적으로 달리해 그라데이션을 준 경우다. 전년도 데이터는 비교를 위한 데이터이므로 금년도 데이터 색보다 연하게 표현할 수 있다.



② 통계표 작성을 위한 데이터 편집

표는 특정 시점의 숫자를 강조할 때 사용하는 표현 방법으로, 읽는 사람이 보면 바로 이해를 할 수 있다는 전제하에 모든 내용을 표현해야 한다. 표도 그래프와 같이 독립, 종속 변수 등이 존재한다. ‘열과 행’ 속 숫자가 그 역할을 한다. 간단한 문장은 짧은 시간 내에 표로 전환할 수 있어야 한다.

숫자형 발표문 해석과 표 구성 방법

제시데이터(메르스 관련 발표자료)

그림 6

중앙메르스관리대책본부가 발표한 메르스 현황입니다. 8일 현재 기준 확진 환자 87명, 사망자 6명, 퇴원 환자 2명, 격리관찰 대상자 2,508명, 격리 해제자 583명, 감염 여부 검사자 76명입니다.

▶ 1단계 데이터가 지닌 중요성 파악

메르스 확진자, 격리자, 퇴원환자, 사망자 등 숫자 하나하나가 지닌 무게감은 매우 크다. 국민의 안전과 직결되기 때문이다. 이처럼 국민에게 알려야 하는 중요 발표 데이터의 경우 지닌 무엇보다 가독성과 표현력에 높은 비중을 두고 제작해야 한다.

💡 분석 P·O·I·N·T

- 자료를 제공하는 기관이 어디인가?(출처 반드시 표기)
- 데이터 수집 시점이 언제인가?(조사시점 반드시 표기)
- 독립변수가 총 몇 개인가?(표로 나타낼 경우 열, 행 개수를 사전 파악하는데 필요)

▶ 2단계 자료 해독

제시자료 독해에 있어서 표에 들어가는 주요 항목, 즉 열과 행에 들어가는 변수를 찾는 것이 우선이다.



중앙메르스관리대책본부가 발표한 메르스 현황입니다.

8일 현재 기준 ① 확진 환자 87명, ② 사망자 6명, ③ 퇴원 환자 2명, ④ 격리관찰 대상자 2,508명, ⑤ 격리 해제자 583명, ⑥ 감염 여부 검사자 76명입니다.

- '중앙메르스관리대책본부'는 주어진 자료를 발표하는 기관명 출처에 포함한다.
- '메르스 현황'은 표 상단에 들어가는 제목이 된다.
- 8일 현재 기준은 데이터를 조사한 시점이다.(표 상단 1행 위치)
- ① 확진 환자 87명, ② 사망자 6명, ③ 퇴원 환자 2명, ④ 격리관찰 대상자 2,508명, ⑤ 격리 해제자 583명, ⑥ 감염 여부 검사자 76명 등 총 6개 변수가 발표 항목이다.
- X축은 독립변수값으로 '사람', Y축은 종속변수값으로 '명'이다.

▶ 3단계(표 구성) 표 만들기-아이콘-배열

1단계 2열 7행으로 구성된 표로 나타낼 수 있다.

그림 8

메르스 현황(8일 현재)	
항목	명
확진 환자	87명
사망자	6명
퇴원 환자	2명
격리관찰 대상	2,508명
격리 해제	583명
감염 여부 검사 중	76명

2단계 단순한 표에 시각적 요소를 넣어 전달력을 높이는 것도 방법이다. 1열 각 행에 적합한 아이콘을 넣어볼 수 있다(픽토그램 : 그래픽 어휘). '확진 환자'는 격리 및 치료라는 이미지, '사망자'는 다른 컬러, '퇴원 환자'는 웃는 모습, '격리관찰 대상'은 돋보기, '격리 해제'는 열린 문, '감염 여부 검사 중'은 현미경과 같은 아이콘으로 각각 나타낼 수 있다.

그림 9

일부 아이콘을 1열에 배열한 모습(시각표 사례)	
메르스 현황	8일 현재
 확진 환자	87명
 사망자	6명
 퇴원 환자	2명
 격리관찰 대상	2,508명
 격리 해제	583명
 감염 여부 검사 중	76명

T-I-P

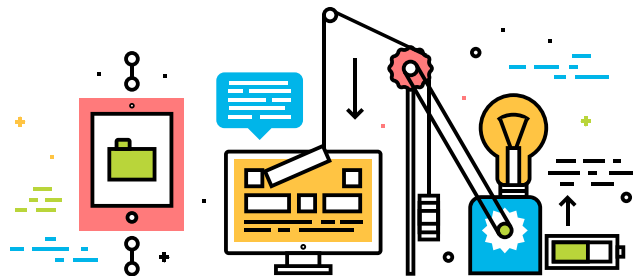
숫자 기반의 데이터를 매일 업데이트해 알려야 하는 경우 시각표로 제작하는 것이 좋으며, 픽토그램을 적절히 활용하는 것도 좋은 방법이다.

▶ 인포그래픽 디자인 제작 사례

메르스 발병 현황을 보여주기 위해 각 항목에 어울리는 픽토그램을 찾아 배치하였으며, 텍스트를 입력할 때 약 모양과 비슷한 도형을 만들어 그 위에 배치하고, 알약의 색으로 현황의 긴급도를 표현했다.



그림 10 (출처: 이수동 & 이해강, 『인포그래픽기획 & 디자인 By 파워포인트』, 예문사)



통그라미, - 자료 수집 어렵지 않아요 -

클릭 한 번이면 나도 통계 전문가! ②



SPSS, SAS, R과 같은 통계 프로그램 패키지는 대부분 자료를 분석하는 기능만 있고 자료를 수집하는 기능은 없다. 이로 인해 자료수집을 할 때는 별도의 프로그램을 사용해야 했다. 하지만 통그라미는 학생들이 자료수집과 통계분석의 두 가지 경험을 모두 할 수 있도록 설계되었을 뿐만 아니라 다른 자료수집 프로그램과 달리 수집한 원자료를 그대로 사용할 수 있는 장점이 있다. 통그라미 초보자들이 설문조사를 할 수 있도록 통그라미의 설문조사의 전체 흐름을 살펴 보고자 한다.

통그라미 접속 및 설문조사

통그라미의 통계분석 기능은 회원 가입 없이 사용할 수 있으나 설문조사 기능은 가입된 회원만 사용할 수 있다. 회원 가입은 크롬에서 '통그라미'를 검색하여 접속하거나 <http://tong.kostat.go.kr/>를 통해 접속한 후 회원 가입할 수 있다.



그림 11 통그라미 메인화면

통그라미에서의 설문조사는 ‘설문지 만들기 - 설문 참여하기 - 설문진도 관리하기’의 순서로 실시하며, 설문 완료된 자료는 ‘자료 관리하기’를 통해 분석에 사용할 자료로 가공할 수 있다.

통그라미는 초등학생부터 사용할 수 있는 프로그램이기 때문에 [그림1]의 ①과 같은 메뉴창 이외에 ②와 같은 바로가기 메뉴가 화면에 같이 나타난다. 메뉴창 또는 바로가기를 클릭하여 설문조사 창으로 이동한다.



설문지 만들기

통그라미를 이용한 설문 조사를 위해 [그림2]의 ‘휴가 때 가고 싶은 해외여행’이란 설문지를 만 들어 보았다.

휴가 때 가고 싶은 해외여행

본 설문은 다가올 휴가 동안에 가고 싶은 해외여행에 대한 의견을 묻는 설문입니다. 작성하신 설문 내용은 통계법 33조(비밀의 보호)에 의거하여 개인의 비밀에 대한 사항은 엄격히 보호됩니다. 설문 종료 후 기재되는 개인정보는 설문조사 응답여부 검증에만 활용되며 통계작성 이외의 목적으로는 사용되지 않습니다.

1. 당신은 다가올 휴가 때 해외 여행을 희망하시나요?
① 예 ② 아니오(8번 문항으로 이동해 주십시오.)

2. 당신이 희망하는 해외여행의 종류는?
① 패키지 여행 ② 자유여행 ③ 테마여행

3. 당신이 희망하는 여행지는?
① 동남아/대만/서남아 ② 중국/홍콩/극동러시아
③ 일본 ④ 남태평양
⑤ 유럽/아프리카 ⑥ 미주/중남미/하와이

4. 당신이 희망하는 여행 기간은? (일)
(※ 예시 : 3박 4일은 4일, 4박 6일은 6일로 기록한다.)

5. 당신이 생각하는 여행의 1인당 금액은? (쇼핑비용 제외)
① 100만원 미만
② 100만원 이상 ~ 200만원 미만
③ 200만원 이상~300만원 미만
④ 300만원 이상~ 400만원 미만
⑤ 400만원 이상

6. 당신이 해외 여행지를 선택할 때 고려하는 사항은? (중복응답가능)
① 기간 ② 비용 ③ 치안
④ 여행목적 ⑤ 음식 ⑥ 쇼핑

7. 당신은 국내여행 보다 해외여행을 선호하시나요?
① 매우 좋음 ② 좋음 ③ 보통 ④ 나쁨 ⑤ 매우 나쁨

8. 당신의 성별은?
① 남 ② 여

9. 당신의 나이는?(만 나이)

♣ 설문에 응해주셔서 감사합니다. ♣

그림 2 | 휴가 때 가고 싶은 해외여행



1단계 설문조사개요 입력하기

통그라미의 설문조사 첫 단계는 설문조사 개요를 입력하는 것이다.

여러 사람들이 사용하는 공간이므로 올바른 언어를 사용해주세요.
타인에게 불쾌감을 주거나, 무의미한 설문은 삭제할 예정입니다.

설문제목

응답기간 : 2019. 11. 03(일) ~ 2019. 12. 03(화)

1 설문제목 : 휴가 때 가고 싶은 해외여행

2 설문설명 : 본 설문은 다가올 휴가 동안에 가고 싶은 해외여행에 대한 의견을 묻는 설문입니다. 작성하신 설문 내용은 통계법 33조(비밀의 보호)에 의거하여 개인의 비밀에 대한 사항은 엄격히 보호됩니다. 설문 종료 후 기재되는 개인정보는 설문조사 응답여부 검증에만 활용되며 통계작성 이외의 목적으로는 사용되지 않습니다.

3 응답기간 : 2019. 11. 03 ~ 2019. 12. 03

4 응답완료메시지 : ✨ 설문에 응해주셔서 감사합니다. ✨

5 공개설정 : ☒ 공개 ☐ 비공개

6 대상자설정 : ☐ 로그인 회원(전체) ☐ 로그인 회원(명부작성) ☐ 참여번호 사용 ☒ 누구나 참여가능(☐ 중복 불가)
☐ 비밀번호 사용(☐ 중복 불가)

7 참여수설정 : 명 (☒ 제한없음)

8 확인

설문추가
표준질문추가
설문대상자 설정
복사보기
설문 삭제
공유하기
복사하기
TOP ^

그림 3 | 설문조사개요

- ① 설문제목을 입력한다.
- ② 설문에 대한 설명을 입력한다.
- ③ 응답 기간을 설정한다. 설정하지 않을 경우 1개월로 기본설정된다.
- ④ 설명조사를 마친 응답자에게 나타날 메시지를 입력한다.
- ⑤ 설문지의 공개 여부를 선택한다.
- ⑥ 설문대상자를 설정한다. 참여번호 사용이나 비밀번호 사용을 이용하여 특정인만 설문에 응하게 할 수도 있다.
- ⑦ 참여자 수를 설정한다.
- ⑧ 확인을 클릭한다.



2단계 설문 문항 만들기

설문조사개요를 입력한 후 설문 문항을 만든다. 통그라미에서 응답유형은 [그림4]와 같이 ‘객관식(단일선택), 객관식(복수선택), 객관식(선호도선택), 주관식’의 4가지로 구분된다.

구분	통그라미 화면
객관식 (단일선택)	<div> <input checked="" type="checkbox"/> 필수항목(필수항목일 경우 체크하세요.) <input type="checkbox"/> 가로배열(문항 가로배열시 체크하세요.) </div> <div> <div>질문</div> <div>당신은 다음을 추가 할 때 해외 여행을 희망하시나요?</div> </div> <div> <div>응답유형</div> <div> <div>객관식(단일선택)</div> <div> <input type="checkbox"/> 답변을 기준으로 질문 이동 </div> </div> <div> <div>예</div> <div>삭제</div> </div> <div> <div>아니오</div> <div>삭제</div> </div> <div> <div>클릭하면 추가</div> <div>또는</div> <div>기타추가</div> </div> </div> <div> <div>질문추가</div> <div>확인</div> </div>
객관식 (복수선택)	<div> <input checked="" type="checkbox"/> 필수항목(필수항목일 경우 체크하세요.) <input type="checkbox"/> 가로배열(문항 가로배열시 체크하세요.) </div> <div> <div>질문</div> <div>당신이 해외 여행자를 선택할 때 고려하는 사항은? (중복응답가능)</div> </div> <div> <div>응답유형</div> <div> <div>객관식(복수선택)</div> <div> <input type="checkbox"/> 기간 <div>삭제</div> </div> <div> <input type="checkbox"/> 비용 <div>삭제</div> </div> <div> <input type="checkbox"/> 치안 <div>삭제</div> </div> <div> <input type="checkbox"/> 여행목적 <div>삭제</div> </div> <div> <input type="checkbox"/> 음식 <div>삭제</div> </div> <div> <input type="checkbox"/> 쇼핑 <div>삭제</div> </div> <div> <div>클릭하면 추가</div> <div>또는</div> <div>기타추가</div> </div> </div> <div> <div>데이터검증</div> </div> </div>
객관식 (선호도선택)	<div> <input checked="" type="checkbox"/> 필수항목(필수항목일 경우 체크하세요.) <input type="checkbox"/> 가로배열(문항 가로배열시 체크하세요.) </div> <div> <div>질문</div> <div>당신은 국내여행 보다 해외여행을 선호하시나요?</div> </div> <div> <div>응답유형</div> <div> <div>객관식(선호도선택)</div> <div> <input type="checkbox"/> 답변을 기준으로 질문 이동 <div>중요우선</div> </div> <div> <div>매우좋음</div> <div></div> </div> <div> <div>좋음</div> <div></div> </div> <div> <div>보통</div> <div></div> </div> <div> <div>나쁨</div> <div></div> </div> <div> <div>매우나쁨</div> <div></div> </div> </div> </div>
주관식	<div> <input checked="" type="checkbox"/> 필수항목(필수항목일 경우 체크하세요.) <input type="checkbox"/> 가로배열(문항 가로배열시 체크하세요.) </div> <div> <div>질문</div> <div>당신이 희망하는 여행 기간은?(※ 예시 : 3박 4일은 4일, 4박 6일은 6일로 기록한다.)</div> </div> <div> <div>응답유형</div> <div> <div>주관식</div> <div>데이터검증</div> </div> </div> <div> <div>질문추가</div> <div>확인</div> </div> <div> <div>데이터검증</div> <div> <div>숫자입력</div> <div> <div>보다큰값입력</div> <div>1</div> </div> </div> </div>

그림 4 | 통그라미의 응답유형

- 객관식(단일선택)은 문항1, 문항2, 문항3, 문항5, 문항8과 같이 보기 중 하나만 선택 가능할 때 사용한다.
- 객관식(복수선택)은 문항6과 같이 보기 중에서 여러 개를 선택할 수 있을 때 사용한다.
- 객관식(선호도 선택)은 문항7과 같이 만족도나 선호도를 조사할 때 사용하며, 선택 시 자동으로 '매우 좋음-좋음-보통-나쁨-매우 나쁨'이 입력되며 수정이 필요할 경우 항목을 클릭한 후 수정할 수 있다.
- 주관식은 문항4, 문항9와 같이 숫자나 문자로 표현된 응답을 수집할 때 사용하며 데이터 검증 등을 통해 입력된 값의 범위를 설정할 수 있다. 예를 들어 여행 기간이 '2일 이상~ 50일 이하'라고 제한할 경우, '숫자입력'과 '보다 큰 값 입력'을 선택한 후 '1'을 입력하고, '숫자입력'과 '같거나 작은 값 입력'을 선택한 후 '50'을 입력한다. 이렇게 설정하였을 때, 만약 응답자가 1 또는 55를 입력한다면 잘못된 값을 입력하였다는 메시지가 나타나게 된다.

그 이외에도 필수항목, 가로배열을 선택할 수 있으며 화살표를 이용하여 항목의 위치를 조절할 수 있다.

3단계 설문 응답에 따른 문항 이동하기

그림 51 문항 이동하기

[그림2]의 설문지에서 문항 1번에서 '아니오'를 선택한 사람은 문항 8번으로 이동하여야 한다. 이 때 사용하는 기능이 '답변을 기준으로 질문 이동'이다.

- ① '답변을 기준으로 질문 이동'을 선택한 후, 문항 이동을 하게 되는 응답인 '아니오'를 선택하면 어떤 문항으로 이동할지 질문이 나타난다.
- ② 이동하고자 하는 문항을 선택하면 설정이 완료된다.
 - ※ 주의해야 할 사항은 문항 1에서 '아니오'를 선택하였을 때, 응답하지 않게 되는 문항 2번에서 문항 7번까지는 필수항목이 아닌 것으로 설정하여야 한다.
- ③ 설문이 완성되면 저장하기를 클릭한다.

4단계 설문지 확인 및 보내기

나의 설문지

설문 제목: [선택] | 설문 상태: [선택] | 검색 | 상세검색

번호	설문 제목	등록일	설문기간	응답수	추천수
4	휴가 때 가고 싶은 해외여행	2019.11.03	2019.11.03 ~ 2019.12.03	9	0
3	새 설문지	2019.11.03	2019.11.03 ~ 2019.11.03	0	0
2	휴가 때 가고 싶은 해외여행	2019.11.03	2019.11.03 ~ 2019.12.03	9	1
1	새 설문지	2019.11.03	2019.11.03 ~ 2019.11.03	0	0

* 모든 설문지
* 모든 설문지

그림 6-1 | 나의 설문지

마이페이지

설문제목: 휴가 때 가고 싶은 해외여행

설문: 본 설문은 휴가 때 가고 싶은 해외여행에 대한 의견을 묻는 설문입니다. 작성하신 설문 내용은 통계로 처리되며, 본 설문은 휴가 때 가고 싶은 해외여행에 대한 의견을 묻는 설문입니다. 설문 종료 후 작성하신 설문은 통계로 처리되며, 설문 종료 후 작성하신 설문은 통계로 처리됩니다.

응답률: 2019.11.03 | 등록자: 통그라미

조회: 0 | 추천수: 0

1. 당일은 다들 휴가 때 해외여행을 떠날까요?

그림 6-2 | 설문수정

저장된 설문지는 [그림6-1]과 같이 ‘마이페이지-나의 설문지’를 통해 확인할 수 있다. 저장된 설문지는 [그림6-2]와 같이 ‘설문수정’을 통해 수정할 수 있다. 설문을 수정한 후에는 반드시 ‘보내기’를 통해 설문조사가 진행되도록 하여야 한다.

설문 참여하기

조사자가 설문지 보내기를 클릭하면 응답자가 설문조사에 참여할 수 있게 된다. ‘설문조사-설문 참여하기’를 클릭하면 [그림7]과 같이 응답할 수 있는 설문의 제목이 나타난다. 이 중 참여하고 싶은 설문 제목을 클릭한 후 조사에 참여한 후 ‘제출하기’를 클릭하면 설문에 대한 응답이 완료된다.

설문 참여하기

설문 제목: [선택] | 설문 상태: [선택] | 검색 | 상세검색

회원: 회원 모두 참여 | 대상자: 대상 회원 참여 | 누구나: 회원, 비회원 모두 참여 | 참여: 참여번호 소유자 참여 | 비밀: 비밀번호 참여

번호	구분	설문 제목	작성자	마감일	참여	상태	결과	SNS공유
793	누구나	휴가 때 가고 싶은 해외여행	통그라미	2019.12.03	0	진행	[보기]	[공유]

그림 7 | 설문 참여하기



설문진도 관리하기

설문조사 현황은 '마이페이지-나의 설문지'를 통해 확인할 수 있다.

번호	설문 제목	등록일	설문기간	문항수	주변수
4	휴가 때 가고 싶은 해외여행	2019.11.03	2019.11.03 ~ 2019.12.03	9	0
3	세 설문지	2019.11.03	2019.11.03 ~ 2019.11.03	0	0
2	휴가 때 가고 싶은 해외여행	2019.11.03	2019.11.03 ~ 2019.12.03	9	1
1	여름 휴가 때 가고 싶은 해외여행	2019.08.28	2019.08.28 ~ 2019.09.28	9	0

그림 8-1 | 나의 설문지 화면

휴가 때 가고 싶은 해외여행

본 설문은 다가를 휴가 동안에 가고 싶은 해외여행에 대한 의견을 묻는 설문입니다. 작성하신 설문 내용은 통계청 33처리실에서 보호에 의거하여 7일간 비공개 처리될 것이며, 보호기간이 끝난 후에는 통계청 33처리실에서 통계작성을 위하여 공개되는 사항입니다.

응답: 1명

1. 당당한 다가를 휴가 때 해외 여행을 희망하십니까? (1명/100%)

답	1	100%
아니요	0	0%

2. 당당한 다가를 여행하는 해외여행의 종류는? (1명/100%)

해외여행	0	0%
자유여행	1	100%
패키지여행	0	0%

그림 8-2 | 설문결과보기

번호	응답현황	완료/미완료	종료일
1	종료됨		2019.11.03 16:34

그림 8-3 | 응답현황

설문 제목 휴가 때 가고 싶은 해외여행

설문 설명 본 설문은 다가를 휴가 동안에 가고 싶은 해외여행에 대한 의견을 묻는 설문입니다. 작성하신 설문 내용은 통계청 33처리실에서 보호에 의거하여 7일간 비공개 처리될 것이며, 보호기간이 끝난 후에는 통계청 33처리실에서 통계작성을 위하여 공개되는 사항입니다.

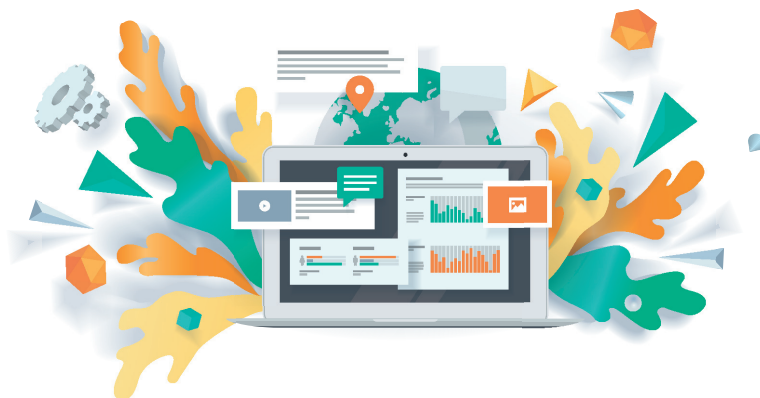
설문기간 2019. 11. 03 ~ 2019. 12. 03

공개일정 ☒ 공개 ☐ 비공개 등록일: 2019. 11. 03

상태 진행 **설문완료처리** **종료일정** 1: 종료

그림 8-4 | 설문완료처리

- ① 관리하고 싶은 설문 제목을 클릭하면 [그림8-1]과 같은 '설문진도관리'를 할 수 있는 화면이 나타난다.
- ② '설문진도관리'를 클릭한 후 설문결과보기, 응답 현황을 클릭하여 [그림8-2], [그림8-3]과 같이 설문 현황을 확인할 수 있다.
- ③ 설문 종료일 이전이라도 [그림8-4]의 '설문완료처리'를 클릭하여 설문을 중단할 수 있다. 중단된 설문조사를 계속할 경우 '설문진행처리'를 클릭하여 설문을 연장할 수 있다.



자료 관리하기

설문조사가 완료되면 수집된 자료 ‘통계분석-자료관리하기’를 통해 자료를 내려받을 수 있다.

그림 9-1 | 자료 관리하기

그림 9-3 | 자료 등록하기

그림 9-2 | 자료 불러오기

- ① [그림9-1]의 자료 관리하기 화면에서 ‘온라인 자료수집’을 선택한 후, ‘자료등록’을 클릭하면 [그림9-2]의 자료 불러오기 화면이 나타난다.
- ② 자료 불러오기 화면에서 ‘나의 설문자료’를 선택한 후 ‘검색’ 버튼을 클릭하면 내가 만든 설문 제목이 나타난다.
- ③ 이 중 내려받고 싶은 설문 제목을 클릭한다.
- ④ 설문 항목이 나타나면 내려받고 싶은 항목을 체크하고 데이터 설명에 간단한 자료명을 입력한 후 확인버튼을 클릭한다.
- ⑤ [그림9-3]의 데이터 편집 화면에서 제외하고 싶은 데이터를 선택한 후 등록완료 버튼을 클릭한다. 데이터 편집에서는 ‘오류검사 설정’을 클릭하여 선택 문항의 응답을 범위를 삭제할 수 있다.

이외의 상세한 설문조사 방법은 통그라미 메인화면의 ‘이용안내 - 이용방법 안내’를 통해 확인할 수 있다.

데이터 경제의 대전환점이 왔다



Interview

한 국 경 제 산 업 연 구 원 경 제 연 구 실 장

김 광 석

“It’s the economy, stupid! (바보야, 문제는 경제야!)”

미국 42대 대통령 빌 클린턴이 대선 때 사용한 캐치프레이즈입니다. 요즘 뉴스를 보면 빅데이터, 4차산업혁명, 사물인터넷, 디지털 트랜스포메이션(digital transformation), AI, 밀레니엄 세대, 블록체인 등 새로운 단어가 쏟아져 나오는 것 같습니다. 이런 용어를 따라가다 보면 결국은 미래 먹거리에 대한 얘기입니다.

이미 많은 전문가들이 데이터가 미래의 석유라는 말을 하며 데이터가 불러올 다양한 사회적 파급 효과에 대해 이야기하고 있습니다. 그래서 오늘은 디지털 경제가 앞으로 어떻게 발전하고 나갈 것인지에 대한 얘기를 해보려고 합니다.

요즘 『경제 읽어 주는 남자』로 어려운 경제에 대해서 쉽고 유익한 정보를 주고 있는 김광석 교수를 만나 디지털 경제에 대한 얘기를 들어 보았습니다.

...▶ **디지털 트랜스포메이션이라는 말이 있는데, 이런 디지털시대는 2000년대 IT혁명과 더불어 시작된 것이 아닌가 하는 생각도 듭니다. 그 시절과 지금의 차이는 무엇이고 왜 이것이 중요한지요.**

디지털 혁명이라는 말은 오래전부터 사용되어 왔는데, 2000년대 인터넷과 PC에 기초한 정보 산업혁명이 일어났습니다. 바로 3차 산업혁명이라고 합니다. 최근 일어나고 있는 변화는 인공지능, 빅데이터, 사물인터넷, 가상증강현실, 블록체인 등의 기술이 산업에 적용되면서 일어나고 있는 4차 산업혁명이라고 볼 수 있겠습니다.

3차 산업혁명부터 아날로그 경제에서 디지털 경제로 전환되는 디지털 트랜스포메이션을 경험하고 있다고 볼 수 있는데요, 최근 4차 산업혁명의 기반기술들이 산업에 확대 적용되면서 혁신의 혁신이 가속화되고 있습니다.

디지털 혁명이라는 말은 오래전부터 사용되어 왔는데,
2000년대 인터넷과 PC에 기초한 정보 산업혁명이 일어났습니다.
바로 3차 산업혁명이라고 합니다. 최근 일어나고 있는 변화는
인공지능, 빅데이터, 사물인터넷, 가상증강현실, 블록체인 등의 기술이 산업에
적용되면서 일어나고 있는 4차 산업혁명이라고 볼 수 있겠습니다.

... 먼저 현재 우리나라 데이터 경제의 경쟁력이 궁금합니다. 그동안 IT 산업이 우리나라가 주도적인
산업으로 성장해왔습니다. 앞으로 4차 산업시대도 우리의 경쟁력은 어떠한지라 생각하시는지요.
5G를 최초로 상용화한 나라, IT 기술 최고 국가, 스마트폰 보급률 1위 국가답게 기존의 전통
산업에서 더욱 더 경쟁력을 높여 나갈 가능성이 크다고 판단됩니다. 다만 어떻게 주어진 능력
과 자원을 활용할 것인지가 중요하겠습니다. 즉 이런 변화에 뒤처지지 않도록 정책 지원과 기
술 개발이 필요합니다. 한동안 휴대전화 시장을 점령했던 노키아의 사례에서 보듯이 그 흐름
을 놓치면 한순간 도태되는 것이 현재의 산업 흐름입니다.

... 데이터 기반 경제 활성화의 조건을 뭐라고 생각하시는지요.

사이버 보안과 규제 완화입니다. 과거 물리적 보안의 시대가 가고 이제 사이버 보안의 시대가
왔습니다. 사이버 보안이 자리잡을 때, 데이터 기반 경제가 자리할 수 있다고 판단됩니다. 또
한, 높은 규제가 데이터 기반 경제로의 이전을 막고 있습니다. 공공 빅데이터의 개방과 기업
및 개인의 빅데이터 공유 과정에서 절대적으로 필요한 것은 규제완화입니다.

... 『디지털 경제지도』란 책에서 세상은 토끼와 거북이의 경주 장소가 산에서 바다로 바뀌는 수준의
디지털 대전환이 이뤄지고 있다고 하셨는데, 땅에서 경주의 법칙과 바다에서 경주의 법칙이 어떻
게 차이가 나는지요.

사람을 접하지 않고 모든 활동이 가능한 ‘비대면화’, 산업 간의 경계가 무너지는 ‘탈경계화
(borderless)’, 개인에 대한 ‘초맞춤화(hypercustomization)’, 제품과 서비스를 완전히 통합하
여 더 나은 가치를 창출하는 ‘서비스화(servitization)’, 데이터의 입력과 동시에 일이 처리되
는 ‘실시간화(real time)’가 그 변화입니다.

이렇듯 대면서비스가 아닌 비대면서비스로, 산업 간의 경계가 명확했다가 이제 산업 간의 경
계가 모호해지고 있습니다. 소비자 만족을 강조하는 경영트렌드에서도 빅데이터를 활용한 초
맞춤화가 가능해지고 있습니다.

제조기업도 단순히 제품을 생산해 전달하는 것이 아니라 이제 서비스의 영역으로 확대·진입하고 있습니다. 생산과 소비가 분리된 경제구조에서 생산과 소비가 실시간으로 이루어집니다. 이제는 이처럼 달라진 환경에 맞는 적절한 대응이 필요한 상황입니다. 환경이 바뀌면 대응도 바뀌어야 합니다.

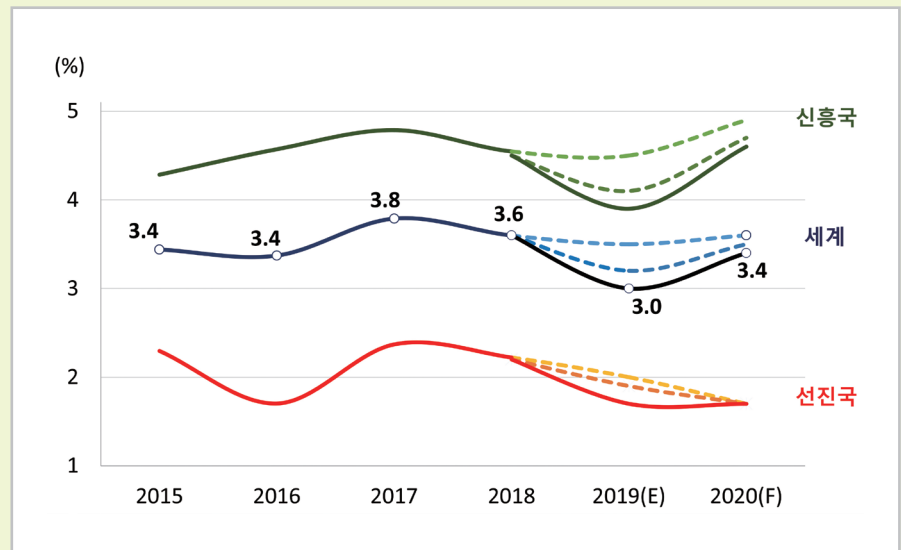
*** 앞에서 말씀하신 변화 중에서 통계와 가장 외닿는 부분이 ‘실시간화’ 인 것 같습니다. 이런 사례가 궁금합니다. 그리고 이렇게 실시간화된다면 데이터 산업에서 가장 요구되는 부분이 어떤 것이라 생각하시는지요.

자율주행차가 대표적인 영역이라고 생각합니다. 자동차들은 수많은 다른 자동차들과 빅데이터를 교환하면서 운영되는 환경에 처합니다(Vehicle to Vehicle, V2V). 그뿐만 아니라 수많은 보행자들과의 통신(Vehicle to Pedestrian, V2P) 및 신호등 등의 인프라들과의 통신(Vehicle to Infrastructure, V2I)이 절대적입니다. 이 모든 것을 V2X(Vehicle to Everything)이라고 합니다. 데이터가 원활히 공유되는 환경은 매우 중요합니다.

*** 이런 변화에서 통계청의 역할도 달라져야 할 것이라고 생각되는데요. 경제학자의 입장에서 통계청에 요구하시는 사항은 어떤 것이 있는지요.

무엇보다 공공빅데이터의 확대 구축과 민간에의 넓은 공유를 선도하는 역할이 필요하다고 판단됩니다. 또 플랫폼 노동자들이 확대되고 있기 때문에, 취업자를 정의하는데 변화가 필요할 것이고, 산업의 경계가 모호해 지고 있기 때문에, 변화된 환경에 맞는 산업의 재정의 등이 필요하다고 봅니다.

2020년 세계경제 전망



자료 : 김광석(2019.10.29) 『한 권으로 먼저 보는 2020년 경제전망』, 이지퍼블리싱.
주 : 2019년 1월과 7월 전망은 각각의 점선으로, 2019년 10월 전망은 실선으로 표시함.

▶▶▶ 아무래도 경제학자를 만나니, 우리나라 경제가 내년에는 어떻게 궁금합니다. 다들 경제가 어렵다고 하는데, 내년 경제를 어떻게 보고 계시는지요.

2020년 한국경제에 있어 가장 주목할 점은 거대한 전환점에 놓여 있는 해라고 말하고 싶습니다. 먼저 소득주도 성장정책이라는 분배정책에서 투자진작이라는 성장정책으로의 방향성 전환이 시작됩니다.

한국경제는 제조업 위기를 맞은 상황이지만, 한일 무역전쟁을 계기로 소재·부품·장비 국산화를 최우선 정책기조로 두는 전환점을 맞이합니다. 산업적으로도 수소경제의 현실화, 소재·부품산업의 집중적 투자, 신재생에너지로의 대전환 등과 같은 탈바꿈된 트렌드들이 준비하게 나타날 것으로 보입니다.

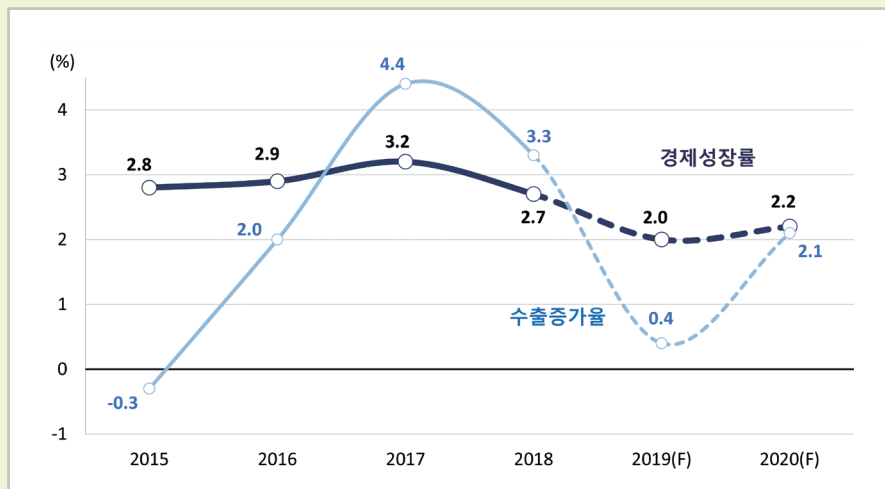
2019년까지 '컨셉'에 머물렀던 디지털 트랜스포메이션은 2020년에는 보다 구체화되기 시작할 것으로 보고 있습니다.

▶▶▶ 내년 경제 위기라는 말이 많은데, 그럼 회복국면에 접어들다고 예상하시는지요.

국제통화기금(IMF), 경제개발협력기구(OECD), 세계은행(World Bank) 등의 세계 주요 경제기구들은 2019년 저점을 뒤로하고, 2020년에는 반등한다고 이야기하고 있습니다. 국내 주요 경제연구기관들도 한국경제가 매우 미미하지만 2020년에 회복국면에 놓여 있다고 말하고 있고요. 이렇게 보는 이유는 기준금리를 인상해 왔던 긴축의 시대가 종결되고, 완화의 시대로 전환될 가능성이 높기 때문입니다. 세계적으로 경기부양에 초점을 두고, 2019년에 고조되었던 긴장감이 2020년 들어 상당 부분 완화될 것으로 보여집니다.

특히 반등하는 신흥국(rebounding emerging)들이 2020년 세계경제의 회복을 견인할 것으로 예상하고 있습니다. 무엇보다도 아시아 신흥국들의 반등세에 주목할 필요가 있습니다. 반면 선진국들은 보호무역조치들과 브렉시트 등의 영향으로 부진한 흐름을 보일 전망입니다.

2020년 한국경제 전망



자료 : 김광석(2019.10.29.) 『한 권으로 먼저 보는 2020년 경제전망』, 이지퍼블리싱.
주 : 수출증가율은 재화의 수출(F.O.B)을 기준으로 함.



...> **끝으로 이런 변화에 대한 우리의 대응 방법에 대해서 한마디 부탁드립니다.**

2020년 한국경제는 2019년 저점에서는 반등하지만, 회복세를 체감할 수 없는 수준의 2.2% 경제성장률을 기록할 것으로 전망됩니다.

하지만 앞에서 말씀드린 것과 같이 2020년은 많은 것들이 바뀌는 대전환점에 있습니다. 세계경제의 기조도, 정부의 재정정책 방향도, 통화당국의 통화정책 결정도 모두 전환되는 시점에 있습니다. 이러한 전환 속에서 변화의 흐름을 적극적으로 관찰하고, 적절한 대응전략을 구사해야 합니다.

정부는 기초체력을 보강하는 정책들에 집중할 필요가 있습니다. ‘경제성장률에 대한 고민’도 필요하지만, ‘잠재성장률에 대한 고민’이 더 중요합니다. 성장잠재력을 확충하기 위해서는 산업구조를 고도화하고 기업들이 경영시스템과 기술을 혁신할 수 있도록 유도하며, 저출산 현상으로 생산연령인구가 감소하는 것을 막는 등의 노력이 요구됩니다.

**김광석 한국경제산업연구원
경제연구실장이 걸어온 길 ...**

- 현대경제연구원 선임연구원 역임
- 삼정KPMG 경제연구원의 수석연구원
- 한국경제산업연구원 경제연구실장
- 한양대학교 겸임교수

저서

- 『경제 읽어주는 남자』
- 『한 권으로 먼저 보는 2019년 경제 전망』
- 『한 권으로 먼저 보는 2020년 경제 전망』
- 『경제 읽어주는 남자의 디지털 경제지도』

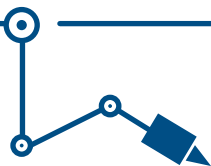


공장에서 나와 현실로 뛰어든 로.봇.

현시대에는 현실 속에서 일을 하는 고성능 로봇이 없을까. 있다. 바로 공장 등 제한된 환경에서 일을 하는 산업용 로봇이다. 로봇이 복잡한 현실사회에서 일을 할 능력이 없으니, 로봇에게 맞는 환경을 만들어 주고 그 속에서 일을 시키는 것이다.

누구나 로봇 산업을 말하는 시대가 왔다. 새 시대에 로봇에 주목해야 하는 이유는 자명하다. 정보화혁명(3차 산업혁명) 이후 지속적으로 발전해 온 소프트웨어 기술은 이미 거대한 시장을 이루며 지구촌의 주력산업으로 자리 잡았다. 그러나 지금까지의 소프트웨어 기술로 만들어낸 가치들은 어디까지나 PC나 스마트폰 같은 ‘정보단말기’ 안에서만 움직였다. 이것만으로도 영상과 소리 등, 정보를 넘어서 엔터테인먼트 요소까지 제공할 수 있었고, 다양한 정보와 감성을 전달하며 세상의 혁명을 이끌었다.

이제 과학기술은 이 벽을 깨고 IT 기술을 무기로 모니터 밖의 세상, 즉 현실세계 역시 소프트웨어로 통제되는 새로운 세상을 만들



어 나가고 있다. 이 상황에서 필수적으로 필요한 것은 IT기술과 현실을 연결하는 매개체, 즉 소프트웨어에 의해 자율적으로 움직이는 기계가 바로 로봇이다.

‘로봇’이라는 단어는 여러 의미를 포함하고 있다. 이 글에서는 자율주행자동차, 무인비행기(드론)를 포함한, 자동화된 기계장치라는 의미로 쓰고 있다. 로봇 없는 새로운 현실에 대응하기 어렵기 때문이다.

시대는 로봇을 요구한다

예를 하나 들어보자. 최근엔 대부분의 차량에 내비게이션 시스템이 붙어있다. 차량의 현재 위치와 목적지의 위치, 운전해서 찾아갈 경로를 보여준다. 이것만으로도 대단한 쓸모가 있으니 누구나 애용하고 있다. 하지만 운전은 내비게이션의 정보를 참고해 사람이 직접 해야만 한다. 즉 자동차라는 ‘현실’ 속의 물건을 움직이려면 어디까지나 사람이 직접 일을 해야 했다.

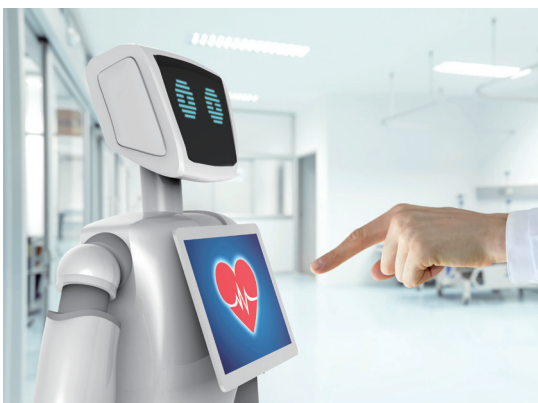
하지만 4차산업혁명 시대가 되면서 이야기가 달라진다. 3차 산업혁명 시대와 4차산업혁명 시대를 구분하는 기준은 사람마다 다르지만 ‘정보화 기기 속 세상과 현실세계가 하나로 합쳐지는 것’으로 보는 시각도 많다. 결국 자동차 산업에서 4차 산업혁명이 도래한다는 말은, 내비게이션의 경로찾기 소프트웨어가 한층 발전하고, 이를 인간이 운전할 필요가 없이 목적지까지 찾아가는 ‘자율주행차’가 시장의 주력으로 부각된다는 말로도 해석이 가능하다.

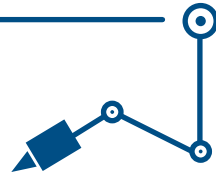
2012년 말 기준 현재 세계 로봇시장 규모는 133억 달러, 이 가운데 제조용 로봇이 약 87억 달러로 전체의 3분의 2에 해당한다. 서비스용 로봇은 46억 달러로 나머지 3분의 1을 차지한다. 하지만 이 중에서 사람을 위해 뭔가 제대로 일을 하는 로봇은 사실 거의 없는 거나 마찬가지이다. 그럼에도 이런 ‘어설픈’ 서비스 로봇조차 관련 시장이 급성장하고 있다. 2012년 세계 서비스용 로봇의 매출은 약 46억4500만 달러, 2007년의 22억3200만 달러에 비하면 2배 이상 시장이 커진 것을 봐도 이런 흐름은 여실히 드러난다.

앞으로 진짜 서비스 로봇 시장, 즉 사람들이 생활하는 현실세계 속에서 로봇의 가치는 점점 더 커질 것이다. 공장용, 산업용 로봇을 넘어서서 현실사회에서 실제로 사람과 함께

움직이는 ‘서비스 로봇’의 세상이 다가오고 있다. 이미 의료용 수술 보조 로봇, 간병 로봇, 서빙용 로봇 같은 첨단 서비스 로봇이 등장하고 있으며, 이런 로봇은 대량생산과 시스템 안전화로 인해 로봇의 가격 역시 점차로 내려갈 것이다. 자동차 한 대씩 구입하는 것처럼 누구나 로봇을 구매하게 되고, 점차 로봇이 인간을 위해 봉사하는 사회가 태동할 것은 불을 보듯 뻔하다.

실제로 해외 선진국들은 이 형태의 산업에 적극적으로 대처하기 시작했다. 일본은 거동이 불편한 환자를 안아 옮기는 로봇, 화장실에 가기 편하게 도와주는 운송로봇 등 제대로 된 서비스로봇 개발에 상당한 공을 들였다. 이미 하체근육이 약한 노인





들의 걸음걸이를 보조하는 하체보조 로봇 할(HAL)의 임대 사업을 벌이고 있을 정도다. 미국과 유럽에서도 하체마비환자를 위한 보행보조 로봇 판매를 시작하는 등 앞다퉈 '쓸모 있는 서비스로봇' 시장이 태동하고 있는 것이다.

로봇과 일자리 시장

서비스로봇의 태동과 동시에 생기는 우려 중 하나는 인공지능으로 무장한 자동화 서비스, 즉 '로봇'이 평범한 직장인들의 '직업안정성'을 빼앗아 갈까 봐 걱정하는 것이다. 사실 이런 우려는 몇 해 전 미국 경제지 '비즈니스인사이더(BI)'의 조사결과 때문에 상당히 증폭되고 있는 듯했다. BI는 향후 20년 안에 로봇(또는 인공지능)이 대체할 가장 유력한 업무'를 발표했는데, 이 조사에 따르면 텔레마케터가 인공지능에 밀려 사라질 일자리 1위로 꼽혔다. 자동응답 시스템을 가진 로봇이 진짜 사람처럼 고객과 대화하며 보증을 팔게 된다는 것. 회계와 감사 업무도 로봇이 대체할 가능성이 높고, 대형 마트의 판매 사원, 컴퓨터 전문용어를 쉽게 풀어 집필하는 '테크니컬라이터'도 20년 안에 없어질 직종이란다. 부동산 중개인과 타이피스트도, 기계기술자와 항공기 조종사, 경제전문가도 로봇 때문에 20년 안에 사라진다고 했다. 하지만 이런 분석만 놓고 '일자리가 사라질 것'이라고 걱정하는 건 어불성설이다. 일부 일자리는 사라지거나 축소되겠지만, 그만큼 로봇 세상이 가지고 올 부가가치 역시 크다. 최근 글로벌 컨설팅회사 맥킨지(McKinsey)는 2030년 기준 제4차 산업혁명이 가져다 줄 지능정보사회에서 발생할 총 경제적 효과는 최대 460조 원에 달할 것으로 예상했다. 인공지능, 그리고 로봇 분야 신규 산업이 활성화되면 이것이 신규매출로 이어지게 되기 때문이다.



먼저 다양한 산업에서 460조 원의 매출 중 무려 44%(약 199조 원)의 비용절감이 일어난다. 단순히 기업마다 돈을 아끼는 것이 아니라, 비용절감을 꾀하기 위해 생기는 새로운 산업의 규모가 이와 같다는 의미다. 또 소비자 삶의 질 향상으로 인한 가치 창출은 38%(174.6조 원), 기술 발달로 인한 신규 매출 발생도 최대 19%(85.4조 원)에 달했다.

자동화된 만큼 일자리는 점점 사라질 거라는 우려도 많다. 맥킨지 역시 또 다른 보고서에서 8억 개의 일자리가 로봇 등장으로 사라질 거라고 했다. 그러나 실제로는 근로시간 전체를 100% 자동화할 수 있는 일자리는 전체 일자리의 약 0.3%에 불과하다고 진단했다. 더구나 자동화율이 10% 미만인 그칠 것으로 예상되는 직업이 무려 97%에 달한다. 이 말의 뜻은, 일부 직업은 자동화로 대체하는 것을 피할 수 없으며, 수많은 일자리가 사라지는 것 역시 피할 수 없지만, 그 대신 로봇산업이 성장하며 새로운 산업과 일자리 역시 큰 폭으로 늘어난다는 의미로 해석이 가능하다.

비슷한 시각은 여러 곳에서 감지된다. 독일 경제정책연구센터(CEPR)는 독일 내에서 로봇이 1대가 늘어날 때 제조업 일자리는 평균 2개가 사라지지만, 제조업 일자리를 줄이는 대신 다른 업종에서 일자리가 늘어나 고용구성에만 변화를 주고 독일 내 총 일자리에는 영향이 거의 없었다고 추정했다.



이는 어찌 보면 당연한 현상이다. 일례로 로봇이 재난현장에 걸어 들어가 공장밸브를 잠그고, 탈출하는 기술을 겨루는 '재난구조로봇경진대회(DRC)'는 2015년 본선을 마치고 2015년 여름 최종 결선에 한국의 KAIST 팀이 우승했다. 현장에서 본 로봇의 성능은 어설피고 답답한 부분이 많았지만 충분한 가능성은 감지됐다. 만약 더더욱 기술이 발전해 전 세계 로봇 전문가들이 고성능 재난구조 로봇을 개발하는데 결국 성공하고, 이 로봇이 능숙한 구조대원 한 사람 몫을 충분히 해 낸다고 치자. 이런 로봇 한 대를 움직이려면 몇 사람이 필요할까. DRC 대회 당시, 주최 측은 대회 당시 너무 많은 인원들이 몰려들까 염려돼 각 참가팀의 인원수를 제한했는데, 40명을 넘어서는 안 됐다. 즉 로봇 한 대가 구조대원 한 사람 몫을 하게 만들기 위해 동원된 석·박사급 엔지니어 숫자가 40명이었다는 말이다. 방시능으로 가득한 위험한 현장에 사람 대신 로봇을 투입할 수 있게 되었을 뿐, 들어가는 인력은 오히려 40배로 늘어났다. 로봇과 인공지능은 정말로 사람의 일자리를 빼앗는 존재일까. 아니면 더 많은 일자리를 만들어 주는 존재일까?

‘로봇 중심 사회’는 반드시 온다

‘팩스 로보티카’라는 말이 있다. 미국의 패권주의를 뜻하는 ‘팩스 아메리카’라는 단어에서 나왔다. 앞으로 사회 전체가 로봇을 중심으로 흘러갈 것이라는 예측이 담겨 있는 말이다. 여러 언론에서도 공공연하게 팩스 로보티카 시대에 대비해야 한다는 말이 흘러나오고 있다.



다들 왜 로봇이 시장의 중심이 될 거라고 여기는 걸까. 그 까닭은 다른 산업에 직접적으로 영향을 미치는 '기반기술'로서 가치가 크기 때문이다. 예를 들어 자동차는 소비제품이지만, 자동차로 다른 자동차나 냉장고를 만들 수는 없다. 하지만 로봇은 이야기가 다른데, 로봇기술을 보유한 기업은, 이를 응용해 같은 산업용 로봇을 만들 수 있고, 군사용 로봇, 가정용 서비스 로봇을 만들 수 있다. 물론 자동차나 냉장고 같은, 로봇이 아닌 전혀 다른 제품도 만들 수 있다. 즉 모든 생산 활동의 근간에 로봇이 들어갈 수 있다는 점을 감안한다면, 그 파급 효과를 무시하기 어렵다. 1차적으로 '가정용 로봇이 몇 대가 팔릴 것인지' 그 수요만 생각해선 안 되는 까닭이다.

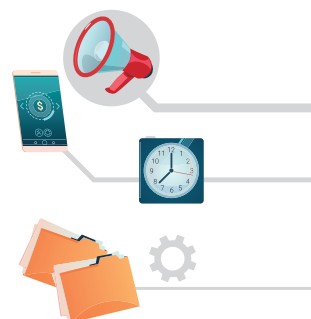
사실 로봇처럼 직접 생산에 관여하진 않지만, 이런 구조를 조금이나마 갖고 있는 분야로 정보기술(IT)을 꼽을 수 있다. 컴퓨터나 스마트기기에 들어갈 반도체나 메모리 구조를 개발할 때도 누구나 컴퓨터를 써서 일을 한다. 그 제품을 판매할 때도, 재고를 조사할 때도, 배송을 할 때도 누구나 IT기기를 써서 일을 한다. 그러니 IT분야 기술은 최소한의 수요를 유지하고, 그래서 시장 장악력을 가질 수 있는 기초산업이다. 더구나 로봇은 IT기기와는 다르게 물리적으로 직접 제품을 생산해 낼 수 있어 그 파급효과는 비교할 수 없이 크다고 볼 수 있다.

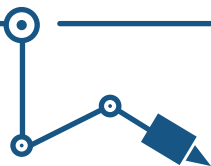
로봇은 모든 지식과 기술이 들어가는 과학기술의 종합상자

로봇이 미래 경제혁신의 견인차로 각광받는 건 현재까지 개발한 거의 모든 지식과 기술이 들어가는 '과학기술의 종합상자' 같은 존재이기 때문이다. 로봇의 몸체를 만들기 위해 필요한 금속 뼈대와 각종 액추에이터, 전압관리, 컴퓨터 시스템 관리 등은 기본이다. 이것들을 하나로 연결해 일괄 조종하는, 소프트웨어를 이용한 제어 기술을 확보해야 한다.

주방 일을 돕는 가사도우미 로봇을 만들려면 관절만 30~40개가 들어가고, 이 관절을 자유롭게 움직이려면 관절 하나마다 1~3개 정도의 액추에이터가 연결돼야 한다. 그러려면 각각의 액추에이터를 제어하기 위해 로봇 속에 전자회로 기관만 수십 장이 필요하다. 여기 연결되는 전선의 숫자는 미처 다 헤아리기도 어렵다. 이 중 전선 하나, 명령어 하나만 잘못돼도 로봇은 맥을 잃고 주저앉는다. 이런 로봇을 생산하고, 판매하고, 유지하려면 지금까지 가전제품이나 컴퓨터를 판매하던 것과는 비교조차 어려운 고도의 생산성과 유지보수 능력도 필요하다. 기업은 훨씬 높은 많은 역량을 요구받고 있다.

이해를 돕기 위해 우리나라 KAIST 연구진이 개발한 인간형 로봇 '휴보'의 내부 구조를 되짚어 보자. 휴보는 상·하반신을 제어하는 시스템이 따로 있다. 사람은 걸어가다가 책상 위에 있는 컵 하나를 집어 들려고 하면 그대로 한 팔을 뻗어 낚아채면 그뿐이다. 하지만 로봇은 전혀 이야기가 다른데, 자칫 중심이 치우쳐 그대로 나동그라질 수 있다. 그러니 상·하체를 제어하는 컴퓨터를 따로 두고, 서로 유기적으로 신호를 주고받으면서 동작을 완성한다. 팔을 뻗으려고 하면 바로 뻗는 게 아니라 하체에 있는 컴퓨터에 신호를 보내고, 상,





하체가 팔을 뻗어도 될 타이밍을 결정한 다음, 하체가 중심을 먼저 이동해 놓은 다음 팔을 뻗어야 한다. 좀 더 세분화해 보면 이야기가 훨씬 복잡하다. 상체가 제어하는 메인 컴퓨터 아래 다시 여러 대의 프로세서가 들어 있고, 하체가 제어하는 메인 컴퓨터 아래도 여러 대의 프로세서가 연결돼 있다. 보통 연산장치 숫자가 20여 개 정도 필요하다.

사실 프로세서가 몇 개가 있느냐 하는 문제는 어디까지나 로봇의 신경망 구조에 대해서만 설명한 것이다. 여기에 소프트웨어를 이용해 일일이 어떻게 제어할지는 순전히 엔지니어 숙제다. 모터의 출력, 하중이 걸리는 위치, 각 부품의 무게, 금속의 피로도까지 모두 감안해야 한다. 재료공학이나 기본적인 물리법칙 등도 완전하게 꿰뚫고 있어야만 가능한 이야기이다. 로봇의 기본 구조를 처음 설계하는 사람 역시 이런 부분을 모두 감안해서 설계해야 한다. 로봇의 하드웨어를 만드는 사람은 어떻게 소프트웨어를 심어서 잘 제어할 수 있을지를 고려해야 하고, 소프트웨어를 만드는 사람들은 이런 하드웨어의 특징을 완전히 이해하고 장점을 최대한 살릴 수 있도록 노력해야 한다.

‘IT’ 산업은 어디까지나 소프트웨어가 중심이다. 물론 하드웨어의 발전도 중요하지만 이는 ‘연산속도가 빨라지면 그만큼 소프트웨어로 더 효율적으로, 더 다양한 일을 할 수 있게 된다’는 것이지, 연산속도 그 자체가 이야기의 핵심은 아니다. 대중에게 가장 큰 영향을 미치는 자동차산업은 어떨까. 반대로 하드웨어가 중심이다. 최근엔 자동화 기술도 많이 활용하고 있지만 자동차의 통제는 어디까지나 사람이 직접 한다. IT산업과 자동차산업은, 이 한 가지만으로도 거대한 산업을 이루기 부족함이 없었다. 이런 점을 고려하면, 로봇이 모든 산업의 중심이 될 수 있다는 시각은 결코 허언처럼 들리지만은 않는다.

한국만의 경쟁력 어떻게 키울까

한국의 로봇시장 구조를 보면 진한 아쉬운 점도 남는다. 한국과학기술기획평가원(KISTEP)은 정기적으로 ‘기술수준평가보고서’를 발간하고 있는데, 세계 과학기술 중 핵심 120개를 추려 주요 5개국(미국, 일본, 유럽, 중국, 한국)의 수준을 고루 비교하는 방식이다.

이 보고서는 각종 첨단기술이 총망라 돼 있지만, 상당부분의 항목에서 미국이 1위를 차지한다. 압도적인 자본력으로 대규모 투자를 이어가고 있는 미국이 세계 맹주의 자리를 지킬 수 있는 숨은 원동력이기도 하다. 하지만 이런 미국도 유럽이나 일본에 뒤지는 분야가 여럿 나오는데, 주로 의료나 헬스케어, 정밀산업 및 서비스 전자기기 등 국가별 특기분야에 집중돼 있다. 처음부터 미국만큼 전 분야에 대규모 자본과 투자를 이어가기 어려우니 자신의 철학을 갖고 특기 분야에 철저히 몰입한 결과다.

하지만 한국은 이도 저도 아닌 상황에 빠져있다. 미국처럼 군사기술의 발전을 기본철학으로 삼고 다양한 분야에 대규모 투자를 할 여력은 없다. 그럼에도 불구하고 투자형태는 거의 유사하다. 군사용 웨어러블 로봇도, 수술용 로봇도, 인간형 로봇기술도 모두 손



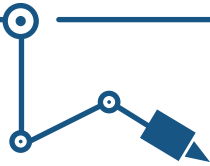


을 뺀다. 그러다 보니 모든 분야에 1위 자리는 일본과 미국, 유럽에 내어주고 많은 분야에서 힘겹게 2위 자리에서 달리기를 계속하고 있는 게 우리나라의 현재 모습이다.

최근 들어 국내 우리나라 로봇기술과 시장은 이미 세계시장에서 크게 인정받고 있다. 거의 대부분의 과학기술분야에서 평균 세계 3~4위권의 기술과 시장을 지닌 것으로 평가받는다. 이런 성과는 정부의 로봇 투자 덕분이기도 하다. 실제로 국내 로봇시장이 2007년 이후 연 평균 약 36.5%의 성장해 왔다는 걸 생각하면 결코 허황된 수치는 아니다. 규모면에서도 괄목할 성장을 해 왔는데, 15여 년 전인 2003년엔 1679억 원 수준이던 국내 로봇시장 규모도 2016년엔 4조4750억 원으로 거의 30배에 달한다.

대분류 코드	대분류명	로봇산업 사업체 수	구성비(%)
100	제조업용 로봇	718	22.8
200	전문서비스용 로봇	300	9.5
300	개인서비스용 로봇	172	5.5
400	로봇 부품 및 부분품	1,001	31.8
500	로봇시스템	461	14.6
600	로봇임베디드	110	3.5
900	로봇서비스	389	12.3
합계		3,151	100.0

출처 : 한국로봇산업협회, 2017년 국내로봇 산업체 현황.



여기서 다시 1년 1년 뒤인 2017년엔 5조5255억 원으로 1년 새 다시 20.2%가 늘어났다. 거의 기하급수적인 성장이다. 세계 로봇시장은 약 15% 정도 성장해 왔다는 걸 생각하면 대단한 고공행진을 계속하고 있다.

첨단 서비스 로봇에 대한 연구 지원 필요

사실 산업용 로봇에 대해서는 세계 1위라는 단어를 써도 크게 이상하지 않다. 2012년 말 기준으로 노동자 1만 명 당 로봇 대수는 396대인데 비해 일본은 322대, 독일은 273대 정도다. 조사 대상 45개국의 평균은 58대라는 점을 비교하면 대단한 수치임에는 틀림없다.

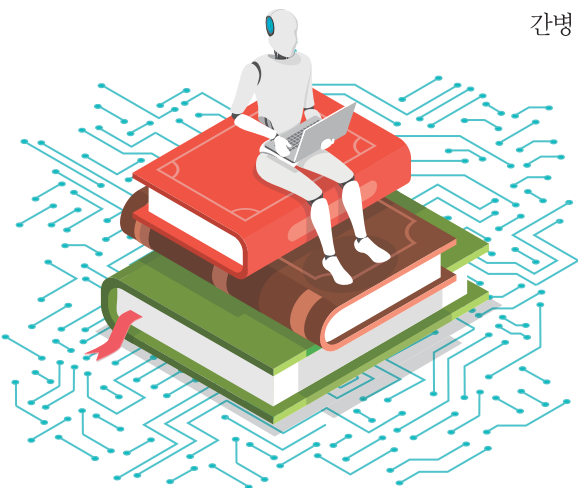
이미 수치상으로 보면 우리나라는 이미 산업용 로봇 분야에서는 미국·일본·독일과 함께 선도그룹에 속한다. 2013년 세계에서 판매된 산업용 로봇 16만2000대 중 70%가 한국·일본·중국·미국·독일 등 5대 로봇 수요 국가에 팔렸다. 국내 로봇 수출액은 점점 증가해 2015년 8159억 원을 기록했다. 2016년엔 여기서 다시 14.4%가 늘어나 9336억 원으로, 2017년엔 1조984원으로 늘어났다.

문제는 로봇의 형태다. 현재 세계 로봇시장의 3분의 2는 제조용 로봇이다. 하지만 1/3을 차지하는 서비스용 로봇의 비율이 급속도로 늘고 있다. 공장에서 일을 시키던 로봇을 변형해, 이제는 앞선 제어소프트웨어 기술과 접목해 실생활에서도 무언가 로봇에게 일을 시키는 것에 집중하고 있다. 2012년 세계 서비스용 로봇의 매출은 약 46억4500만 달러, 2007년의 22억3200만 달러에 비하면 2배 이상 시장이 커진 것은 이런 흐름을 잘 대변한다.

국내를 비롯해 세계적으로 판매 중인 서비스 로봇의 수준도 문제다. 대부분은 그리 높은 수준의 기술이 필요 없는 '청소용 로봇' 같은 것들이 상당수다. 그 이외에는 과거의 IT 기술이 기반한 오락용 제품이 대부분이다. 이런 서비스 로봇은 사실상 로봇이 아니라 컴퓨터인데, 로봇 모양을 한 컴퓨터가 대부분 화면과 음성으로 서비스를 하는 형태다. 미래에 로봇이 인간과 함께 살아가는 삶, 로봇 혁명에 대비하려면 이런 구조와 철학을 갖고는 다소 취약점이 있다. 실제로 로봇의 서비스 산업을 고려해야 한다. 의료용 수술 로봇,

간병 로봇, 서빙용 로봇 같은 첨단 서비스 로봇의 비중이 점점 높아질 것을 고려하면, 관련분야 연구를 집중적으로 지원해 나갈 필요가 있다.

현시대에, 분명 한국의 로봇산업은 현재 변화의 중심에 있다. 한국로봇기술이 선진국을 모방해 이 정도 수준에 도달한 것은 우리나라 산업발전 형태와도 닮아있다. 앞으로 한국이 이런 역량을 살려 코앞으로 다가온 서비스로봇 시대를 본격적으로 대비해 나가야 할 때다. 우리만의 독자적인 철학이 무엇인지 고민하고, 거기에 적합한 로봇산업 생태계를 구축할 필요도 있다. 이는 우리나라 정부와 연구기관, 그리고 기업들이 함께 고민해야 할 숙제다.



빅데이터 시대... 가장 핫한 대회가 생겼다!

통계빅데이터 분석 활용대회



요즘 빅데이터가 단연 화제이다. 이런 시대에 대응하고자 통계청은 「통계빅데이터센터」를 설립하여 산·학·연 및 일반 국민이 활용할 수 있도록 통계빅데이터를 제공하고 있다고 한다.

통계빅데이터센터는 통계빅데이터를 연계·활용할 수 있는 데이터 분석 플랫폼으로 엄격한 보안체계가 갖춰진 공간에서 자유롭게 데이터를 분석할 수 있도록 지원하는 곳이다. 통계청 정부혁신 차원으로 2018년 말부터 본격 운영을 시작하여 현재 서울·부산·대전에 설치되어 있다(<http://data.kostat.go.kr>). 여기서 빅데이터를 활용하고 이런 대회에 응모도 할 수 있다니 이것이 바로 핑 먹고 알 먹고가 아닌지 모르겠다. 몰랐다면 이제 알고 가자.

‘제1회 통계빅데이터 분석 활용대회’ 어떻게 진행되었나

- **목적** → 정부혁신의 일환으로 통계빅데이터센터 자료 분석 기회를 제공하여 데이터 기반의 의사결정 문화 확산 및 통계빅데이터센터 이용 활성화
- **주제** → 데이터로 통하는 세상
공공데이터와 민간데이터를 융·복합하여 사회·경제적 현상 분석 및 정책 의사결정 모델 발굴, 창업 아이디어 등 제안
- **지원자격** → 개인 또는 팀(3명 이하)의 형태로 누구나 참여 가능
*외국인의 경우 개인 참가는 어려우며, 한국인을 포함한 팀으로 참가 가능
- **대회기간** → '19. 6. 3(월)~8. 2(금)
- **활용데이터** → 통계빅데이터센터 제공자료, 공공·민간데이터
 - 타출처의 데이터와 연계 분석 가능
 - 통계빅데이터센터 제공자료
 - 행정통계자료(통계등록부(10종), 행정통계(8종)), 민간자료(SKT 빅데이터)

행정통계자료

부문	자료명	시계열
통계등록부	통계기업등록부DB(사업자 기준)	2010-2017년
	통계기업등록부DB(동일인 기준)	2010-2017년
	사업장 기초DB(4대보험)	2010-2017년
	사업장 기초DB(법인)	2010-2017년
	등록센서스 인구DB	2015-2017년
	등록센서스 가구DB	2015-2017년
	등록센서스 주택DB	2015-2017년
	농업DB	2015-2017년
	임업DB	2015-2017년
	어업DB	2015-2017년
행정통계 (승인통계)	귀농·귀촌인통계(귀농)	2013-2017년
	귀농·귀촌인통계(귀여)	2013-2017년
	귀농·귀촌인통계(귀촌)	2013-2017년
	신혼부부통계(전체)	2015-2017년
	주택소유통계(개인)	2012-2017년
	주택소유통계(가구)	2015-2017년
	중장년층행정통계(개인, 가구)	2016-2017년
	영리법인기업체행정통계(10차)	2015-2017년

민간자료(SKT)

부문	자료명	시계열
인구·가구	성·연령별 유동인구	2015.01.-2018.12.
	시간대별 유동인구	2015.01.-2018.12.
	요일별 유동인구	2015.01.-2018.12.
	성·연령별 유입인구	2017.07.-2018.06.
	성·연령별 주거인구	2016-2017년
	성·연령별 직장인구	2016-2017년
	가구소득	2016-2017년
경제·사회	성·연령별 매출	2015.01.-2018.06.
	업종별 매출액(현대카드)	2015.01.-2018.06.
	1000대 상권 정보	2018년
	영업 중 업소	2017.07.-2018.06.
건물·지가	표준공시지가	2016-2017년
	주택(단독/다세대/기타)	2016-2017년
	공동주택정보	2016-2017년
	건축물정보	2017.07.-2018.06.
차량	기초정보 인구집중유발시설	2016.-2018. 분기
	도로별 차량통행량	2017년
기초정보	시·군·구별 차량등록대수	2017년
	행정통코드	2018년 분기
기타	법정통코드	2018년 분기
	SNS데이터	2017년

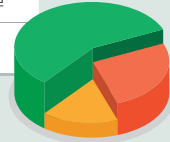
* 행정통계자료와 마이크로데이터·민간자료 간 개별자료 연계를 위한 키값 없음



‘제1회 통계빅데이터 분석 활용대회’ 결과는

학생, 연구원 등 다양한 분야의 141개 팀이 참여하여 최종 61건의 보고서를 제출하였다. 제출된 보고서 중 주제의 참신성, 데이터 활용성, 분석 수준 등을 고려하여 전문가 심사위원회를 통해 최종 7편의 수상작을 선정하였다.

상급	제목	소속	성명
최우수	골목식당을 중심으로 살펴본 방송 프로그램의 골목 시장 활성화 효과 평가	성균관대학교 경제학과	문상균
우수	신도시 타당성 요인분석	경기대학교 응용통계학과	김동현, 최종욱, 박소담
	지역의 사회구조적 특성이 빈집 형성에 미친 영향에 관한 패널 분석	한국형사 정책연구원	한민경
장려	서울특별시 버스 노선 혼잡도 예측을 통한 다람쥐 버스 신규 노선 제안	이화여자대학교 통계학과	고명지, 유채연, 박정민
	빅데이터를 활용한 사업부지 맞춤형 컨설팅	전주시청	김성숙, 이강식, 고미숙
	#축제 거기 어때?	충청지방통계청	김영아, 정예슬, 장선일
	92년생, 82년생, 72년생, 62년생, 52년생 김지영	뉴스1	서영빈



최·우·수·상

골목식당을 중심으로 살펴본 방송 프로그램의 골목 시장 활성화 효과 평가

- **분석취지** → 방송 프로그램을 통한 특정 가게의 문제점 개선과 방송노출이 주변 지역의 상권 활성화에 미치는 효과를 분석하고자 한다.
- **활용데이터** → 통계청 통계빅데이터센터에서 제공하는 민간자료 중 성·연령별 유동 인구 자료와 업종별 매출 자료를 이용하였다.
- **분석방법** → 식당의 위치와 유동인구 혹은 매출을 측정한 관측 지점 간의 거리를 이용하여 지역을 구분한 후 방송 시작 전후 시점을 비교하는 이중차분법 연구설계를 통해 ‘골목식당’ 방송의 효과를 추정하였다.
- **분석결과** → 추정결과 ‘골목식당’ 방송은 유동인구와 음식업 매출을 증가시키는 효과가 나타났고 특히 유동인구는 장기적(6개월 후)으로 유지되는 경향을 보여주었다. 다만, 영향범위는 200m 수준으로 제한적이며, 매출의 효과는 음식업 이외 상권 활성화로 이어지는 것은 아닌 것으로 나타났다.



우·수·상

신도시 타당성 요인분석

- **분석취지** ... 기존 1, 2기 신도시 개발 효과를 분석하여 최근 발표한 3기 신도시 개발의 방향성에 대해 제언하고자 한다.
- **활용데이터** ... 통계청 통계빅데이터센터에서 제공하는 통계기업등록부, 민간데이터(성·연령별 유동인구, 매출액, 상권정보) 및 이용자 보유자료를 사용하였다.
- **분석방법** ... 신도시 성공기준은 계획인구와 현 거주인구의 일치 또는 상회 여부로 간주하였고, 교통, 상권, 자금능력을 영향요인으로 채택했다. 분석에는 주성분분석과 표준화 거리를 이용한 k-근방 분류를 사용했다.
- **분석결과** ... 분석을 통해 1기 신도시에 비해 2기 신도시는 성공한 신도시의 요인인 교통, 상권, 자금능력 부분에서 아직 상대적으로 부족함을 확인했다. 3기 신도시 목표 완수를 위해 부천(중동)신도시 연구, 지하철 신설 및 확충 또는 기존노선 연장 등을 제언한다.

우·수·상

지역의 사회구조적 특성이 빈집 형성에 미친 영향에 관한 패널분석



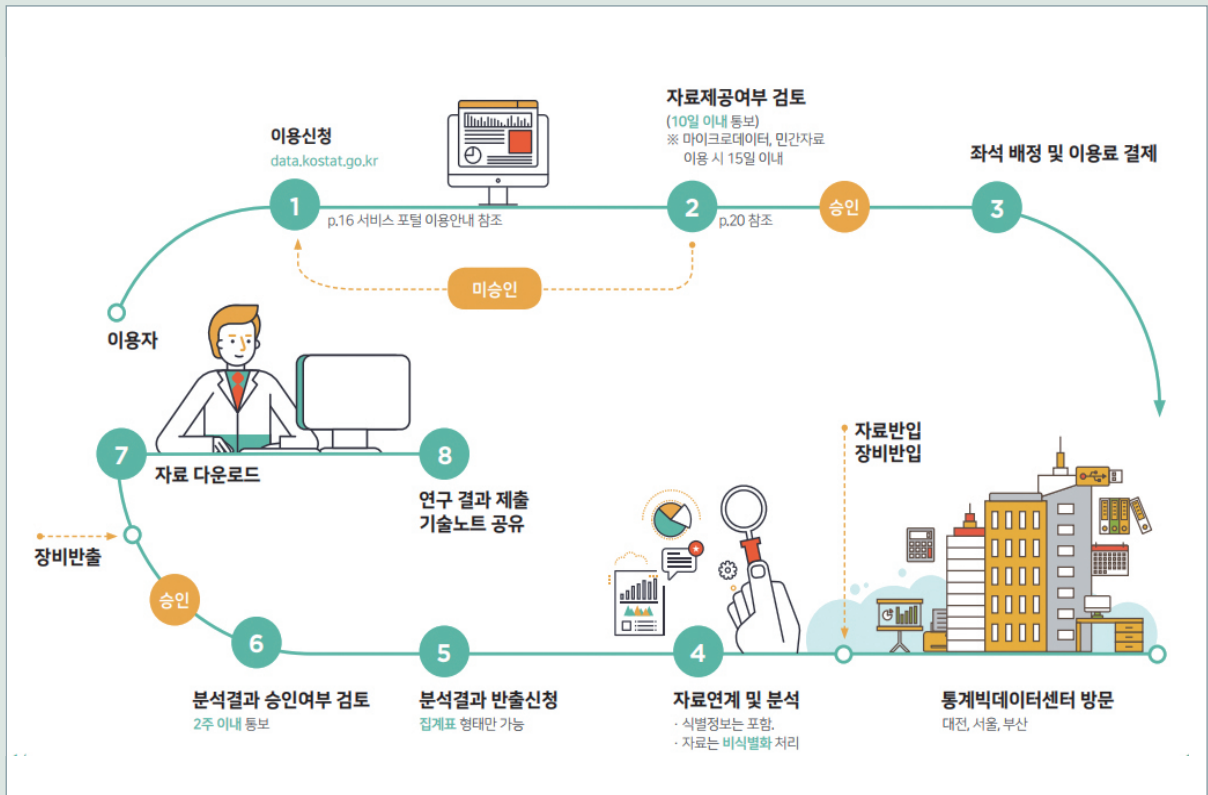
- **분석취지** → 농촌 뿐 아니라 도시 빈집도 증가하는 사유는 저출산·고령화 이외에 사회구조적 취약성이 영향 미치는지 분석하고자 한다.
- **활용데이터** → 통계청 통계빅데이터센터에서 제공하는 행정통계자료 중 등록센서스 인구DB, 가구DB, 주택DB 및 통계기업등록부DB를 활용하였다.
- **분석방법** → 읍면동 다년도 패널데이터를 구축하고 고정효과에 따른 포아송 모델로 분석하였다.
- **분석결과** → 빈집 발생 및 증가 현상은 인구구조 특성 외에도 주택 상황* 및 경제적 상황, 생활편의 수준이 복합적으로 작용한 결과로 분석되었다. * 아파트수, 오피스텔수, 취약주택수, 연면적, 건축연도, 노후주택수, 신축주택수

‘통계빅데이터센터’를 이용하려면

통계빅데이터센터는 이용자가 공공데이터를 편리하게 이용하고 민간자료 및 외부 반입 자료와 연계·융합이 가능하도록 구축된 데이터 플랫폼이다.

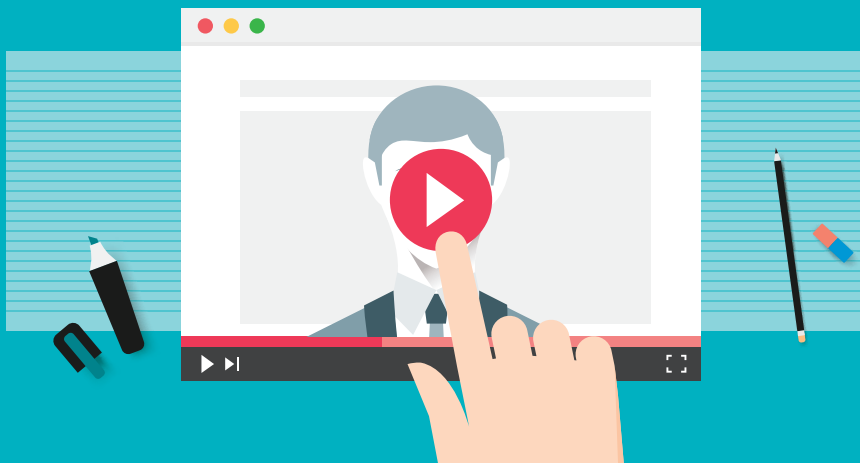


• 이용 방법



영상 편집...

이것만 알아도 나도 유튜버



사진이 아닌 영상시대가 오다

얼마 전 보람튜브가 유튜브 수익으로 수십억이 공개되어 한창 이슈였습니다. 보람튜브는 보람이라는 한 아이의 브이로그를 올리는 채널입니다. 브이로그는 들어보셨나요? 브이로그(Vlog)는 비디오 로그, 즉 영상기록으로 최근 유튜브를 통해서 텍스트가 아닌 비디오로 소통하면서 생긴 신조어입니다. 스마트폰으로 쉽게 영상을 찍고 유튜브를 통해 쉽게 영상을 올려 주변에 알리고 수익도 창출되면서 많은 이들이 이 브이로그에 관심을 갖게 되었습니다.

하지만 막상 브이로그를 시작하려면 영상을 찍고 편집하는 기술이 어려워서 쉽게 접근이 어렵습니다. 가족여행의 추억을 영상으로 이리저리 꾸며 보려 해도 영상 편집이라는 벽에 가로막혀 그냥 포기하는 경우가 많습니다. 그래서 영상편집에 대해 전혀 몰라도 우리가 늘 갖고 있는 휴대전화의 인공지능으로 쉽게 영상을 편집하고 추억을 만들 수 있는 방법을 준비했습니다.

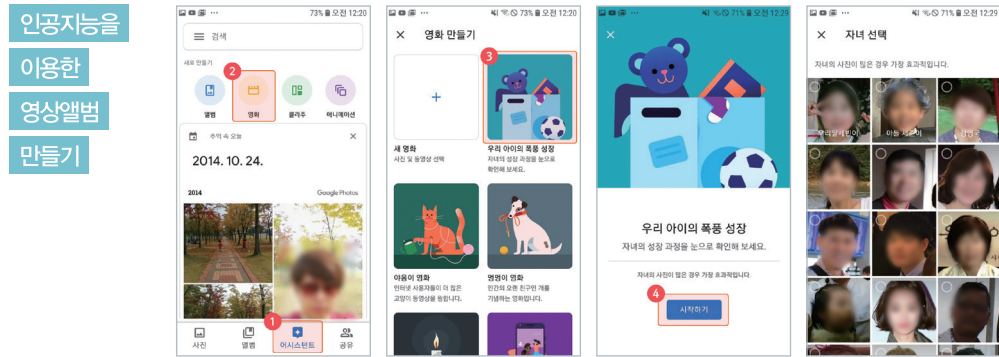
초보자를 위한 인공지능 영상편집 방법

영상을 편집하는 단계는 크게 컷편집, 음악, 자막 및 효과 이렇게 세가지로 나뉘 볼 수 있습니다. 휴대전화를 이용한 영상편집은 갖고 계신 휴대전화의 운영체제(OS)에 따라 방법에 차이가 있기 때문에 먼저 휴대전화의 OS를 확인하셔야 합니다.

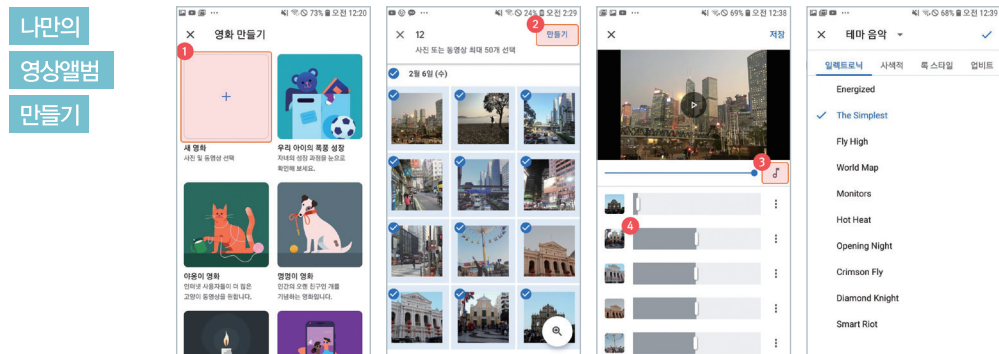


내 폰이 안드로이드OS라면 구글 포토 앱을 활용해 보자

갖고 계신 휴대전화가 안드로이드 OS라면 구글 포토 앱을 통해서 영상을 손쉽게 편집할 수 있습니다.
(만약 휴대전화에 Google 포토 앱이 설치가 되어있지 않다면 구글플레이에서 앱을 먼저 설치해주세요.)



→ 앱을 실행하고 ❶ [어시스턴트] 탭을 선택합니다. 상단 '새로만들기'에서 ❷ [영화] 메뉴를 탭합니다. '영화만들기'에는 다양한 메뉴들이 있습니다. 이곳에서 인공지능으로 '우리 아이의 폭풍성장' '행복한 한 해' '셀카영화' 등 영상을 만들 수 있습니다. ❸ [우리 아이의 폭풍 성장]을 탭해서 아이의 성장앨범을 만들어보겠습니다. ❹ [시작하기]를 탭합니다. 마지막 화면으로 아이의 사진을 선택하면 끝. 인공지능이 자동으로 영상앨범을 만들어줍니다. 쉽죠!



인공지능을 이용해서 영상을 만들면 내가 원하는 사진이나 영상이 누락되기도 합니다. 내가 원하는 영상으로만 만드는 '나만의 영상'은 어떻게 만들까요?

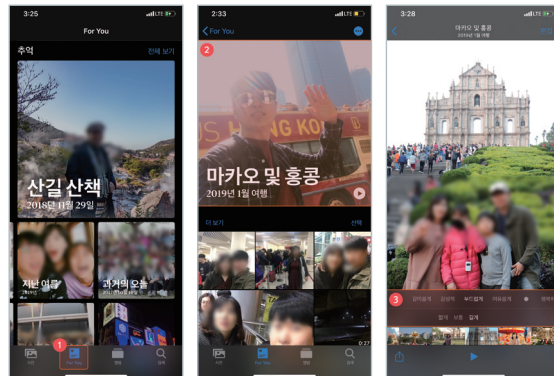
→ ❶ [새 영화]를 선택하여 저만의 영상을 만들어 보겠습니다. 앨범에서 영상으로 만들 클립 또는 사진들을 선택합니다. 우측 상단에 ❷ [만들기]를 탭하면 편집화면으로 넘어갑니다. 여기에서 음악을 바꾸고 싶으면 우측의 ❸ [음악] 아이콘을 선택해서 원하는 음악을 선택합니다. 소장하고 있는 음악이 없다면 [테마음악]에서 원하는 음악을 선택합니다. ❹ 각 클립 또는 사진의 시간 길이를 조절하여 컷 편집을 합니다. 편집이 끝나면 저장을 누르면 완성이 됩니다.

내 폰이 iOS라면 사진앱을 이용해 보자

iOS의 경우는 안드로이드와 달리 직접 영상을 만드는 것이 아니라 인공지능이 만들어 놓은 초안을 수정하는 방식입니다. 마치 나를 위한 집사가 “주인님 이 영상앨범이 맘에 드십니까?”라고 물어오는 느낌입니다. 음악과 컷 편집에서도 원하는 감정의 느낌을 선택하게 되어있습니다. 「감미롭게」「감성적」「여유롭게」등으로 말입니다. 그럼 iOS의 기본 영상편집을 해보겠습니다.

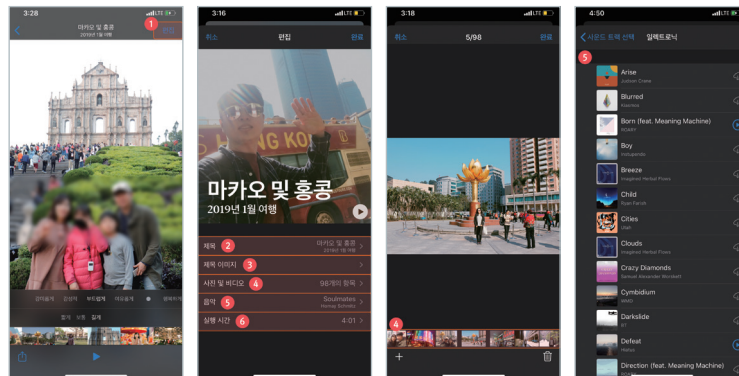


인공지능을 이용한 기본 편집



→ [사진]앱의 ❶ [For You]를 탭합니다. '추억'이라는 카테고리에 iOS의 인공지능이 만들어 놓은 영상앨범이 있습니다. 영상을 선택하면 완성된 초안이 있습니다. ❷ [화면]을 탭하면 기본화면이 보여집니다. 여기에서 전체적인 영상의 ❸ 느낌을 선택할 수 있습니다. 그리고 영상의 길이를 설정할 수 있습니다.

기본 영상앨범 만들기



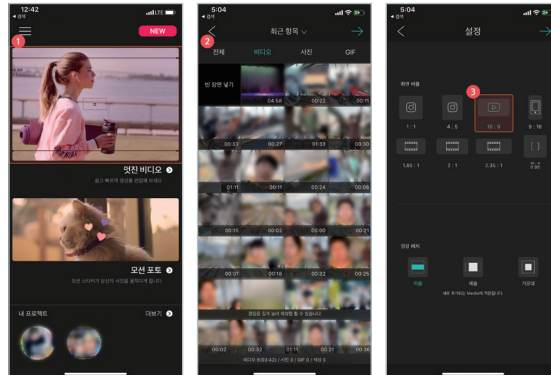
자, 이제 본격적으로 편집을 해보도록 하겠습니다.

→ 우측 상단에 ❶ [편집]을 탭하면 편집화면으로 넘어갑니다. ❷ 제목과 부제목을 입력할 수 있습니다. ❸ 제목이미지와 대표이미지를 선택합니다. ❹ 사진 및 비디오에서 영상에 사용할 클립들을 선택합니다. ❺ 음악을 탭하면 음악을 선택할 수 있습니다. 사운드 트랙을 선택해서 음악을 고르고 원하는 음악을 다운받아서 영상에 사용할 수 있습니다. 마지막으로 ❻ 시간길이를 조절하고 저장하면 됩니다. 쉽죠!

중급자를 위한 인공지능 영상편집

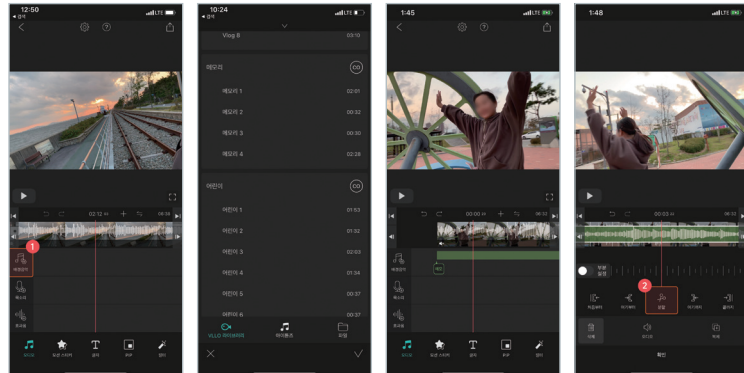


기본으로 제공되는 영상편집에서 만들어진 영상이 조금 심심한가요? 더 전문적인 영상을 하고 싶으신가요? 좀 더 시간을 투자해서라도 자막을 넣고 영상전환(트랜지션)도 포함된 영상으로 만들고 싶으신가요? 그렇다면 영상편집 VLLO앱을 이용해보세요. (*앱스토어 또는 구글스토어에서 VLLO앱을 다운받습니다.)



→ 앱을 실행하면 ❶ [멋진비디오], [모션포토] 메뉴가 있습니다. 영상 편집을 위해 ❶ [멋진비디오]를 선택합니다. 최근 항목에서 ❷ 영상에 사용할 클립 또는 사진을 선택합니다. 설정화면에서 화면비율과 영상배치를 선택합니다. 유튜브에 올릴 영상이라면 ❸ [16:9]를 추천합니다. 처음에도 언급했듯이 영상편집의 기본은 1. 컷 편집, 2. 음악, 3. 자막 및 효과입니다. 영상편집의 하나의 팁을 드리자면 음악을 먼저 깔고 음악에 맞게 컷을 편집하는 것입니다.

배경음 컷 편집



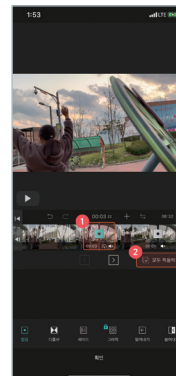
그럼 먼저 배경음을 선택해볼까요?

→ ❶ [배경음] 아이콘을 탭하고 리스트에서 음악을 선택합니다. 이때 대부분의 음악은 유료이고 메모리와 어린이가 카테 고리는 무료로 사용이 가능합니다. 무료 음악도 충분히 쓸만합니다. (원하는 음원을 가지고 있다면 파일에서 불러와 사용할 수 있습니다.) 음악을 불러오면 초록색 막대 바가 스트리라인 아래 생겼습니다. 이제 ❷ 음악을 들으면서 박자에 맞춰 컷을 분할하고 다음 컷을 이어줍니다. (TIP 이때 박자를 찾는 것이 어렵다면 배경음의 초록색 막대를 탭해 보세요. 소리의 높낮이가 시각적으로 표시되어 박자 찾기가 쉬워요)

트랜지션

(장면전환)

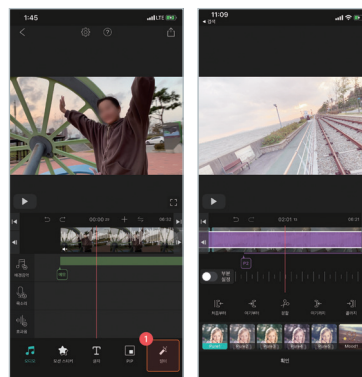
→ 음악에 맞춰 컷 편집이 완성됐다면 필요한 부분에 트랜지션 효과를 넣어줍니다. 컷과 컷 사이에 ❶ 네모 상자가 있는데 이 상자를 선택하면 디졸브 페이드 밀어내기 등의 트랜지션을 넣을 수 있습니다. (tip.하나의 트랜지션을 선택하고 ❷ [모두 적용하기]를 눌러주세요. 모든 컷에 동일한 트랜지션을 넣을 수 있어요)



필터

(컬러그레이딩)

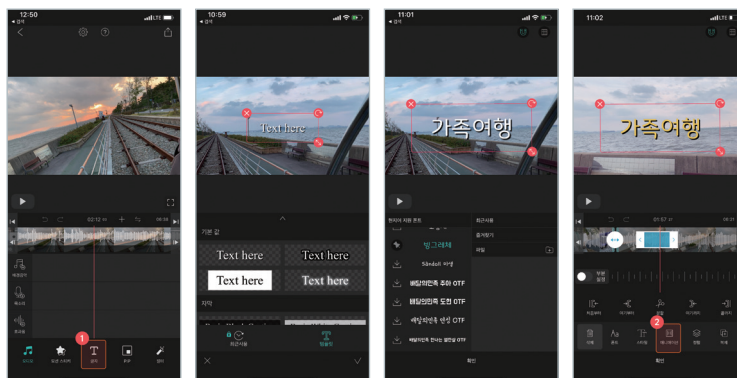
→ 트랜지션까지 넣었다면 필터를 넣어보겠습니다. 필터를 넣으면 영상의 색상을 바꿔주면서 영화 같은 분위기를 낼 수 있습니다. 메인편집화면 하단의 ❶ [필터]탭을 선택합니다. 필터에서 다양한 필터 중 마음에 드는 필터를 선택합니다. 이제, 마지막으로 필요한 곳에 자막과 효과를 넣어보겠습니다.



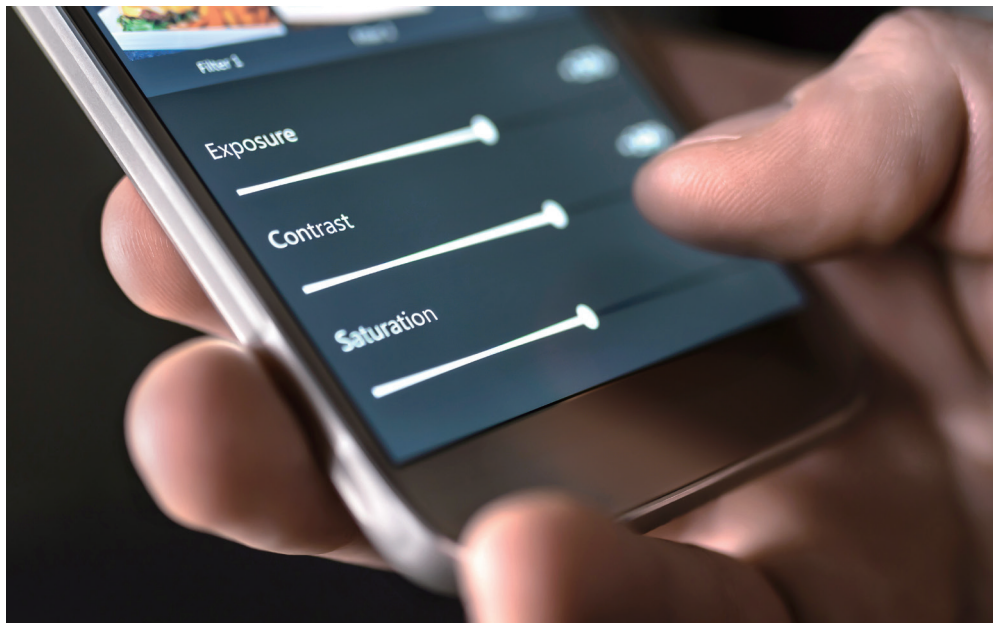
자막 및

모션스티커

(특수효과)



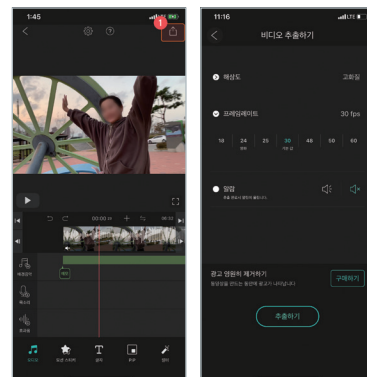
→ 자막과 모션스티커의 효과를 넣으면 영상이 훨씬 역동적이고 풍부해집니다. 먼저 첫 화면의 시작에 제목 글자를 넣어 보겠습니다. 메인 편집화면에서 ❶ [글자]를 선택합니다. 타임라인의 왼쪽에 [글자+]를 선택하고 스타일을 정합니다. 화면 위에 텍스트박스가 나타납니다. 텍스트박스를 더블탭 하거나 아래의 '글자' 메뉴를 탭하면 글자를 편집할 수 있습니다. 그리고 폰트와 스타일을 꾸며주세요. ❷ [애니메이션] 효과까지 넣으면 더 역동적인 영상이 만들어집니다.



영상

내보내기

→ 자, 이제 영상편집이 완성되었습니다. 해상도와 프레임 레이트를 확인 후 [추출하기]를 선택하면 영상파일로 추출이 됩니다. (**TIP** 프레임레이트는 초당 프레임 수를 뜻합니다. 기본 30fps를 쓰고 영화에서는 24fps를 많이 사용합니다) VLO앱을 이용하여 함께 영상편집을 해보았습니다. 처음엔 어려울 수도 있지만 몇 번 해보면 멋진 영상을 여러분도 만들 수 있습니다.



마치며

OS별로 기본으로 제공하는 영상편집과 VLO앱을 이용한 영상편집을 알아보았습니다. 어떤가요? 어렵지 않게 영상을 만들 수 있죠? 이제 이렇게 만들어진 한 편의 추억을 혹은 하루의 일상을 SNS를 통해 많은 사람들과 소중한 추억을 공유할 수 있습니다. 어서 시작해보세요!



당신이 잠든 사이, 몸은 만들어진다.

2007년 세계보건기구(WHO) 산하 국제암연구소는 ‘야간근무’를 2A급 발암물질로 정했다. 납 화합물, DDT 살충제, 디젤엔진 배출물 등의 고악한 화합물과 같은 등급이다. 바꾸어 말하면 수면이 그만큼 중요하다는 것이다. 달콤한 잠을 자는 동안 몸은 스스로 치유하고 성장한다. 낮 동안은 나도 모르게 거북이처럼 목을 앞으로 쏙 빼며 일했지만, 잠자는 동안만큼은 자연스러운 중립자세를 찾아야 한다. 우리는 한평생 3분의 1이라는 긴 시간을 잠자는 데 사용하므로, 좋은 수면 자세가 땀 흘려 운동을 하는 것만큼이나 근골격계 건강에 중요하다.

의료계는 알고 있는 질병에 따라 수면자세를 알려준다.

역류성 식도염 환자는 왼쪽으로 누워 자면 좋다. 내부 장기가 일렬로 되면 역류하는 산의 양이 줄어서 속 쓰림이 완화된다. 위산이 역류되



소리치는 자는 범인이
아니라 피해자이다

— 물리치료사, 다이애나 리(Diana Lee)





지 않게 베개도 높은 것을 추천한다. 고혈압환자는 심장에 여유 공간을 만들어주는 오른쪽이 좋다. 하지만 임신부는 이 자세가 별로 좋지 않다. 특히 임신 말기에는 태아로 가는 혈류에 제한될 수 있기 때문이다.

천장을 보고 똑바로 누워 자면 얼굴의 주름과 반점이 덜 생겨 피부에 좋다. 하지만 옆으로 자는 사람보다 코골이와 수면 무호흡증이 걸릴 위험이 2배나 높다. 코골이 때문에 수면무호흡증에 시달리면 옆으로 누워서 자는 것을 추천한다. 그러나 질병이란 근본 치료를 하는 것이 원칙이다. 치료는 하지 않고 편향된 수면자세로 증상조절에만 매달리면 궁극적으로는 근골격계 질환 하나를 더 얻는다.

잘못된 수면자세는 몸을 서서히 망가뜨린다. 몇십 년에 걸쳐 몸의 정렬을 무너뜨리기 때문에 누구도 수면자세가 범인이라고는 눈치채지 못한다. 더군다나 수면자세라고 해봤자 머리를 똑바로 또는 옆으로 한다

든지, 손바닥을 위로 또는 아래로 하다든지 등의 사소한 동작이기 때문에 설마

이것 때문에 몸이 아플 것이라고는 생각하지 않는다. 그러나 고혈압, 고지혈증, 고혈당 등의 중요한 내과 질환이 생활 습관병인 것처럼 근골격계 질환 역시 생활 습관에서 오는 경우가 많다.

그렇다면 요즘 사람들은 어떤 모습으로 자고 있을까? 한 병원에서 최근 1개월간 성인남녀 142명을 대상으로 ‘평균 수면자세’를 조사한 결과, 24%(34명)만이 ‘차렷’형 자세로 잠을 잔다고 응답했다. ‘옆으로 누워 잔다’ 21%(30명), ‘옆치락뒤치락’ 19%(27명), ‘(태아처럼 웅크린)새우잠’ 18%(25명), ‘옆으로 누워 하반신만 비틀어진 자세’ 12%(17명), ‘옆드린 자세’ 6%(9명) 순이었다. 결국 ‘차렷’형 자세로 잠을 자는 성인은 10명 중 2명에 불과했다.

척추 정렬이 올바른 사람은 똑바로 누운 자세에서 특별한 불편함을 느끼지 않지만, 퇴행성척추질환이나 요통환자는 똑바로 자면 통증이 있다. 척추관협착증이나 허리디스크가 있는 사람도 똑바로 누우면 척추관이 좁아지고 하반신 신경을 압박해 다리가 저릴 수 있다. 옆으로 누워 자면 척추관이 넓어지면서 일단 통증은 준다. 반대로 척추전만이 있으면 옆드려 다리를 구부린 채 자면 편하게 느껴진다. 이처럼 우리가 느끼는 편한 자세라는 것도 통증을 감소시키는 자세이지 반드시 바른 수면자세는 아니라는 것이다.

근골격계 질환에 도움이 되는 바른 수면자세는 ‘차렷’형 자세와 비슷한 ‘해부학적인 자세’이다. ‘해부학적인 자세(anatomical position)’라는 생소한 말은 의학공부를 시작하는 사람이라면 누구나 가장 처음 배우는 용어이다. 그만큼 기본자세이며 모든 의학적 기술의 기준점이 되는 중립자세이다. 발, 손바닥을 앞으로 향하여 똑바로 선 자세로서, 공식 해부학적 명명법으로 정해진 각종 구조 또는 부위, 혹은 방향을 기술하는데 참조하는 자세로 사용된다.

‘해부학적 자세’는 모든 관절이 펴진 상태로 근육은 이완과 수축이 없는 편안한 상태다. ‘차렷’ 자세와 다른 점은 손바닥이 몸통 옆이 아니라 앞을 보고 있다는 점이다. 손바닥이 앞으로 향하면 팔꿈치 아래 있는 두 개의 뼈가 꼬임이 없어지고 날개뼈 안쪽 근육도 긴장이 풀린다. 자연스럽게 가슴이 펴져 어깨가 앞으로 말리지 않는다.



육중한 우리 몸이 중력을 이기고 걸어 다니기 위해서 신체는 체중을 고르게 분산되는 구조로 되어야 한다. 그것이 바로 척추의 S라인이다. 'S라인'이란 목과 허리뼈가 C자형을 이루고 가운데 있는 흉추가 역 C자를 만들어 내는 긴 S자를 말한다. 25개의 척추뼈가 만든 S라인은 마치 큰 스프링처럼 충격을 흡수해 걷거나 움직일 때 체중을 골고루 분산시킨다. 이러한 척추 만곡이 없다면 체중이 고스란히 뼈로 전달되고 근골격계의 퇴행성 변화를 지금보다 훨씬 빨리 겪게 될 것이다. 바른 수면자세를 강조하는 이유 중의 하나도 이런 척추 만곡이 수면 중에 파괴되는 것을 최대한 막자는 것이다.

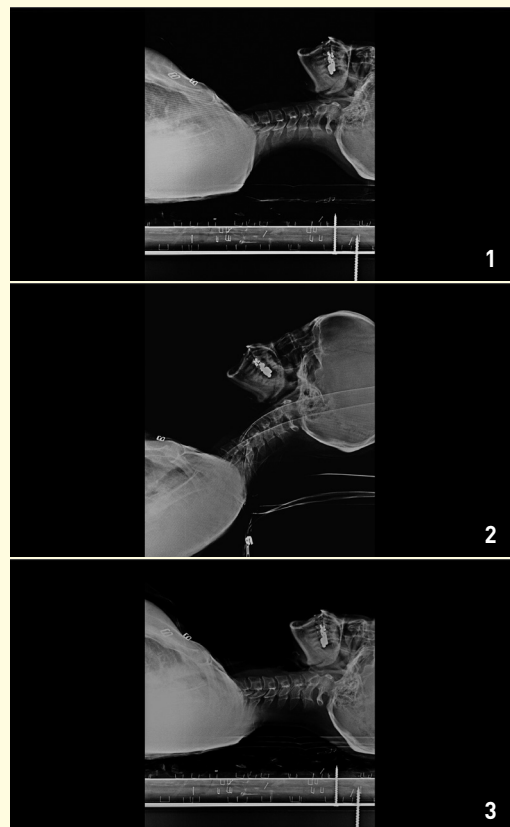
그래서 목 베개가 필요하다. 척추 중 가장 힘없고 가냘픈 목이 자면서도 C자 정렬을 유지하려면 뒤에서 받쳐주는 구조물이 필요하기 때문이다. 그것이 없다고 당장 목이 내려앉는 것은 아니지만 오랜 시간이 지나면 목 주위의 근육과 뼈의 정렬은 변하게 된다.

헬스트레이너 김 관장은 베개가 몇 개인지 모른다. 몇 해 전 베트남 여행에서 사온 라텍스 베개부터 친구가 추천한 호주에서 직수입한 경추베개, 유명한 정형외과의사가 디자인했다는 고가의 베개까지 베개만 해도 이불장 한 가득이다. 또 베개를 만드는 소재 또한 고급스럽다. 솜이나 메밀보다는 라텍스, 메모리폼, 폴리우레탄, 편백나무 조각 같은 것을 사용한다. 모두가 김 관장만큼은 아니라도 베개 한 개로 몇 년을 버틴 옛날과 달리 요즘은 베개 두세 개는 보통이다. 그만큼 베개에 대한 관심도 많고 베개시장도 활성화되어있다.

고침단명(高枕短命)이라는 옛말만 믿고 아예 베개 없이 자는 사람도 있다. 자는 동안 목을 받쳐주지 않으면 멀쩡했던 C자형 목이 수년에 걸쳐 일자가 된다. 일자 목이란 엑스선 촬영에서는 뼈가 단 순히 일직선 배열을 한 것으로만 보이지만, 목 주변의 근육이나 인대가 비정상적으로 짧아지거나 늘어졌다는 것도 알아채야 한다. 목은 좁은 공간에 머리, 얼굴 그리고 팔로 가는 혈관과 신경이 촘촘히 밀집해 있다. 머리와 몸통을 연결시켜주는 목을 고속도로로 비교해보자. 8차선 고속도로쯤은 돼야 신경과 혈관이 무리 없이 지나 다닌다고 가정하면 목 근육이 비정상적으로 수축과 경직이 되어 2차선으로 변해 버리면 목을 지나는 수많은 혈관과 신경이 딱딱해진 근육 사이에 끼이게 된다. 이런 현상은 불필요한 두통, 안구통, 어깨 결림, 팔 저림 등의 비전형적인 증상에 시달리게 한다.

목이 C자가 아니면 목을 둘러싼 근육의 균형이 무너져 있고 신경이 활성화되어 잠자는 사이에도 근육이 계속 일을 한다는 것의 의미한다. 이때 잘 다치는 근육은 날개뼈와 경추를 연결하는 견갑근과 목빗근, 승모근 등이다. 목과 어깨근육은 연결되어 있어서 병도 항상 친구처럼 따라 다닌다. 그래서 목이 아픈 사람이 나중에 어깨도 나빠져 있다.

베개를 사용하지 않는 사람과 반대로 높은 베개가 편하다고 자꾸 베개를 높이게 되면 목은 일자목보다 더 위험한 역 C자가 된다. 역 C자는 목 디스크로 발전할 수 있다. 베개와 목 건강, 수면건강에 대한 논문도 많이 나오고 있지만, 어떤 베개를 선택해야 하는지는



1 경추베개를 사용했을 때의 목뼈의 모양.

2 높은 베개를 사용했을 때의 목의 모양.

3 베개를 사용하지 않을 때의 목의 모양.



의사도 잘 모른다. 확실한 것은 베개란 모름지기 머리가 아닌 목을 잘 받쳐야 한다는 것이다.

나는 목이 앞으로 빠지지지는 않았지만 일자목이다. 현미 채식으로 유명한 대체의학을 하는 분이 알려준 '베개 없는 수면'을 13년 했다. 그래서 지금은 자는 동안 목이 C자를 이루도록 긴 원통모양의 베개로 목을 받친다. 뒤통수의 볼록한 정도에 따라서 개인차는 있지만 목 베개의 높이는 5~6cm면 적당하다. 옆으로 누울 때는 똑바로 때보다 조금 높게 7~8cm다. 이 높이가 척추가 옆에서 봐서 일직선을 이루게 할 것이다. 경추베개가 없으면 수건 두 장을 탄탄하게 돌돌 말아 고정시켜서 사용한다. 자다보면 목 꺼지는 경추베개보다는 직접 만든 수건 목 베개가 훨씬 낫다. 그러나 일자나 역 C자 목이 아주 심한 사람이 바로 경추베개로 바꾸는 것은 위험하다. 목 근육치료를 충분히 한 뒤 차츰 베개를 바꾸어야 한다.

목을 받쳐주었으면 다음은 손이다. 대부분의 사람들은 손등이 위로 보이게 두 손을 살포시 모으고 배 위에 올려놓고 잔다. 손바닥을 위로하는 것과 손등을 위로하는 것이 동전의 앞과 뒤처럼 비슷한 의미가 아니다. 손은 날개뼈를 통해서 목뼈로 연결되어있기 때문에 손을 어떤 모양으로 하느냐에 따라서 어깨와 목이 긴장하는 정도는 달라진다. 목이 앞으로 빠지고 어깨가 말린다는 소리를 듣는다면 부처님처럼 손바닥을 앞으로 보이는 자세만 해도 많이 좋아진다. 팔짱끼는 자세나 손을 배 위에 모으고 자는 자세는 어깨를 앞으로 점점 더 굽게 만든다.

적당한 높이의 목베개를 하고 두 팔을 손바닥을 위로해서 몸통 옆에 가지런히 두면 그것만으로도 가슴이 펴지고 어깨가 제자리를 찾는다. 스쿼트나 벤치프레스만이 근육을 멋지게 만드는 것은 아니다. 근육은 섬세하고 정확한 동작으로 잘 쉬게만 해줘도 스스로 멋지게 회복할 것이다.

사람에 따라서는 새우처럼 웅크려 자야 잠이 든다. 마치 어머니 뱃속에 있는 것처럼 아늑하고 일시적으로 허리통증도 덜하다. 그러나 그렇게만 자면 요통은 악순환이 된다. 어쩔 수 없이 밤새도록 새우잠을 잔 날에는 밤새 구겨져 있던 '장요근'을 꼭 늘려줘야 한다. 장요근은 뱃속 안에서 허리에서 대퇴부를 연결하는 길고 큰 근육이다. 한참동안 쏘그리고 앉았다가 갑자기 일어설 때 허리가 잘 안 펴지는 것도 이

근육이 수축과 이완을 빨리 하지 못해서이다. 사무실 의자에 앉아 있다가 잠잘 때조차 허리를 꼬부리고 새우잠을 자면 장요근은 주구장창 수축만 한다. 그래서 다른 건 몰라도 새우 잠자는 사람은 장요근 스트레칭은 꼭 해줘야 한다. 한쪽 다리를 앞으로 세우고 반대편 다리를 뒤로 죽 편다. 한손은 구부린 무릎 위에 다른 한 손은 바닥을 짚는다. 뒤로 뻗 다리의 골반이 늘어나도록 앞으로 지긋이 스트레칭한다.

오래 사는 세상이다. 하지만 아직까지 우리나라는 아프면서 오래 사는 세상이다. 재밌게 살다가 마지막에는 온 뼈마디가 징그럽게 아파서 요양원에만 틀어박혀 살아야 하는 거야. 누가 행복한 인생이라고 하겠는가? 끝이 좋아야 좋은 법이다. 그래서 나는 '성공한 노화란 근육을 100세까지 88하게 쓰는 것이다.'라고 생각한다. 통증 없이 건강하게 오래 살기 위해서는 근육운동, 단백질공급과 더불어 무심했던 수면자세도 꼼꼼하게 점검해야 한다. 몸은 당신이 잠드는 동안에도 아주 천천히 만들어진다.



직장인도 작가가 될 수 있나요?



좋아하는 걸 쓰면 됩니다

“선생님, 책은 어떻게 쓰나요?”

작가 10년차를 넘어서면서 직장(초등학교) 동료들에게 가장 많이 받은 질문이다. 그 질문을 다시 곱씹어 보면 묻는 사람의 의도를 쉽게 알 수 있다. 내게 질문을 한 사람들 역시 책을 쓰고 싶은 마음이 있다는 말. 그리고 그들도 작가가 되고 싶다는 말이기도 하다. 그들에게 난 이렇게 답한다.

“좋아하는 걸 쓰면 됩니다.”

그들에게 한 나의 이 대답은 나를 작가로 만든 말이기도 하다.

초등학교 6년, 그리고 중·고등학교 6년, 이렇게 총 12년의 학창시절을 보내면서 난 글짓기대회나 독후감대회에서 단 한 번도 상을 받은 적이 없다. 아니, 더 엄밀히 말하자면 원고지 세 장을 넘기는 일이 너무도 어려웠다. 항상 의무적으로 써야하는 글짓기 대회의 최저 분량은 원고지 3장. 난 대부분의 세 번째 장을 ‘니다.’ 혹은 ‘다.’로 마무리했다. 이런 내가 작가가 되었다는 건 토종닭이 남쪽으로 날아가는 기러기 무리와 하늘에서 합류하여 남쪽으로 날아가는 일과 비슷할지도 모른다.

내게 책을 처음 권한 사람은 다름 아닌 직장 선배교사였다.(지금 이 글을 읽는 여러분들에게는 내가 그 선배교사일지도 모른다.) 그는 이미 여섯 권이나 되는 책을 집필한 중견 작가였다. 신간이 나오자 그 선배교사는 같은 6학년 교사들에게 자신이 집필한 책을 한 권씩 나눠주었다. 그리고 한마디 덕담처럼 우리 모두에게 책을 써보라고 권했다.

그날 책을 쓰고 싶다는 생각을 한 사람은 ‘모인 사람 전부’였고, 그 생각을 실행으로 옮긴 사람은 ‘나’뿐이었다. 여기서 실행으로 옮겼다는 말은 원고를 쓰고 그 원고를 출판사에 투고한 사람을 뜻한다.

그 선배교사의 말 중에 내게 가장 다가온 말은 이렇다.

“자신이 가장 좋아하는 것을 조사해서 모은 후, 그것을 그럴싸하게 엮으면 책이 됩니다.”

이 말은 지금 작가가 된지 십년이 지난 나에게도 항상 유효하다.

작가가 좋아해서 쓴 책은 매력이 있다

그날 난 선배교사의 말을 듣고 내가 가장 좋아하는 것을 떠올려 보았다. 가장 먼저 떠오른 건 ‘음악’이었다. 그리고 그와 거의 동시에 ‘수학’도 떠올랐다. 다른 것들은 애써 꺼내야 되는 것들이라 일단 제외시켰다. 음악과 수학 중에 고민하던 난, 방향을 수학 쪽으로 결정했다. 우리 반 아이들이 가장 싫어하는 게 수학이었기 때문이다.

주제가 정해지자 내가 해야 할 일이 별 고민 없이 떠올랐다. 일단 내가 쓰고 싶은 수학 관련 서적을 조사하는 것이다. ‘아이들이 싫어하는 수학을 재미있게 이야기로 풀어낸 책’을 찾기 위해 대형 서점을 찾았다. 그 미션을 수행하기 위해 제목과 목차, 그리고 작가의 말을 먼저 읽었다. 그리고 취지에 맞으면 앞의 다섯 장 정도는 꼭 읽었다. 이 과정을 반복하고 또 반복하다가 내가 그토록 찾던 책을 찾았다. 바로 김성수 작가의 『피타고라스 구출작전』이라는 책이었다. 그 전까지의 살펴본 책들이 대개 다섯 장이나 여섯 장에서 내 손을 벗어났다면 이 책은 끝장을 넘길 때까지 내 손을 붙드는 강력한 힘이 있었다. 다 읽고 나서 내 마음에 든 생각은 ‘나도 이런 책 한 권 쓰고 싶다.’였다.

난 책을 구입해서 곧장 집으로 갔다. 그리고 집에서 다시 찬찬히 읽으며 분석을 시작했다.

이 책에는 내가 좋아하는 ‘수학’, 아이들에게 필요한 ‘재미와 흥미가 있는 수학’이 모두 들어 있었다. 그리고 이야기 자체도 흥미진진해서 수학을 싫어하는 아이들도 이야기로만 충분히 읽을 만한 책이었다. 특히 이 책에 나오는 수학 문제들은 모두 알쏭달쏭하면서 풀고 싶은 것들로 이루어져 있었는데 이 부분이 정말 매력적인 책이었다.

좋아하는 것이 주는 선물, 그리고 피드백

이제 내가 다음에 해야 할 일이 명확해졌다. 일 년간 6학년 아이들에게 수학을 가르치며 재미있는 문제를 만들기도 하고 여러 방법을 찾아 변형을 시키는 작업을 했다. 아이들은 문제 자체에 빠져들어 재미를 느꼈고, 난 그런 문제들만 모아 따로 정리해두었다.

문제가 어느 정도 준비가 된 후, 난 원고 집필에 들어갔다. 글을 쓰며 내가 놀란 것은 원고가 이상하리만큼 술술 써진다는 것이었다. 원고지 세 장을 쓰기 어려웠던 내가 단지 나이가 들었다는 이유로 글을 잘 쓰게 된 것일까라는 생각도 들었지만 이유는 다른 데 있었다. 바로 내가 좋아하는 것에 대해 글을 썼기 때문이다.

내가 좋아하는 ‘수학’ 그리고 아이들이 싫어하는 ‘수학’을 이야기로 재미있게 담아낸 원고가 한 달 정도 걸려서 완성이 되었다. 그리고 퇴고를 하는 데 두 달을 사용했다. 여러 차례 퇴고를 하고 나는 아내에게 제일 먼저 원고를 건넸다.(사실 아내 말고 내 원고를 성의 있게 읽어 줄 사람이 그 당시에는 없었다. 사실 지금도 내 책의 첫 독자는 항상 아내다.)

원고를 읽은 아내는 나에게 형식적인 말인지는 모르겠지만 재미있다고 평가해주었다. 그 말에 자신감을 얻은 난, 다음날 바로 원고를 출판사에 보냈다.



출판사에 투고하는 방법

출판사에 투고하는 방법은 생각보다 간단하다. 한 마디로 원고를 워드프로세서로 작성해서 출판사의 홈페이지 투고란에 올리면 된다. 몇몇 출판사들은 출력본을 요구하기도 하기 때문에 등기를 보내야 하는 경우도 있지만 대개 웹상에서 투고가 가능하다. 나의 경우는 원하는 출판사가 모두 홈페이지에서 가능했기에 홈페이지 투고란을 이용하기로 했다. 내가 선택한 출판사는 『피타고라스 구출작전』을 출판한 ‘주니어김영사’와 어린이 책을 많이 출판하고 있던 ‘웅진주니어’였다.

원고를 업로드하고 하루, 이를 시간이 흘러갔다.

‘언제쯤 연락이 오는 걸까? 혹시 선택받지 못한 원고는 연락조차 하지 않는 걸까?’

다양한 생각들이 한 주간 내 머리 속을 가득 메웠다. (사실 대부분의 출판사들은 원고를 받으면 빠르면 하루, 늦어도 한두 달 안에는 답신을 준다. 물론 원고가 채택되느냐 그렇지 않느냐와는 별개의 일이다. 경험상 빠른 연락은 출판 제의가 들어오고 한 달 이상 지난 경우는 대개 출판의 의사가 없음을 알리는 메일이다.) 그로부터 며칠이 지난 어느 날, 원고를 보낸 지 일주일이 지나기 전에 주니어김영사에서 연락이 왔다. (내가 웅진주니어보다 더 바랐던 출판사였다.) 내가 처음 써서 보낸 원고가 놀랍게 채택이 된 것이다.



책이 만들어지기까지

그 이후로 일 년 가까운 기간 동안 초고는 수정에 수정을 거쳐 점점 책다운 모습을 갖추어 갔다. (일반적으로 퇴고 기간은 작가의 역량에 따라 다른데 초보 작가나 직장인 작가의 경우, 대개 육 개월에서 일 년까지 꽤 오래 걸린다. 하지만 중견 작가의 경우는 한두 달 만에 퇴고가 끝나는 경우도 많다.) 마치 커다란 조각상을 만드는 과정처럼 틀을 잡고 점점 섬세하게 다듬어 나가는 퇴고의 과정은 예술작품을 만드는 과정과 비슷하다. 이 과정에서 초보 작가들은 대개 편집자와 소통을 하며 글을 다듬는 것뿐 아니라 출판과정에 대한 많은 것들을 배우게 된다. 이런 과정을 통해 초보 작가는 중견 작가로 발전하게 된다.

내 원고가 거의 완성되자 편집자는 그림 작가를 섭외했다. 그림 작가의 선택은 글 작가와 출판사가 협의하에 결정하게 된다. 한동안 그림 작가와 편집자, 그리고 나, 이렇게 삼자간의 교류가 이루어지고 그림 작업이 끝나자 일러스트 작업에 들어갔다. 상대적으로 일러스트 작업기간은 짧았다. 두 주 정도 되었던 걸로 기억한다. 사실 이진 경우에 따라 다르긴 하다.

처음 문제를 만드는 것부터 글을 쓰고, 원고를 완성해서 그림에 일러스트까지 완성하는 데 걸린 시간은 거의 2년 정도 걸렸다. 이렇게 해서 나의 첫 책인 『12개의 황금열쇠』가 출판되었다. 이 책은 독자들에게 반응이 상당히 좋았는데 그 덕에 대만으로 번역되어 수출까지 되었다.

처음 책 쓰기에 도전하는 당신에게

직장을 가진 사람들은 전업 작가처럼 매년 여러 권의 책을 집필하기는 어렵다. (실제로 지인 중 필력이 좋은 전업 작가 몇 분은 일 년에 수십 권의 책을 쓰기도 한다.) 하지만 일 년에 한 권이나 여러 해에 걸쳐 한 권의 책을 출판하는 것은 그렇게 어려운 일이 아니다. 앞서서도 말했듯이 자기가 좋아하는 것을 조사해서 잘 정리한 후, 그럴싸하게 엮으면 정말 책이 된다.

처음 책 쓰기에 도전하는 경우, 두려움이 먼저 앞서기 마련이다. 하지만 그 두려움은 잘 써야 한다는 강박관념 때문에 오는 것이기에 잘 써야 한다는 생각만 버리면 간단하게 해결된다. 다시 말해서 자신감을 가지고 시작하라는 말이다. 그리고 이렇게 자신감을 가지고 시작했다고 하더라도 막상 모니터 앞에



서 자판에 손을 올리면 눈앞이 캄캄해지면서 아무것도 생각이 나지 않을 수 있다. 그럴 땐, 좋은 음악을 듣거나 재미있는 영화 한 편을 보고 그냥 푹 쉬어도 좋다. 하지만 이런 날이 계속되면 책 쓰기와는 점점 거리가 멀어진다.

만일 일주일 정도 이렇게 글을 쓰지 못한 채 흘러가면 다른 대책을 적용할 시기가 도래한 것이다. 이번엔 처음처럼 컴퓨터 앞에 앉아 멍하니 모니터만 보고 있더라도 최소 한 시간은 글을 쓰려고 노력하며 버텨야 한다. 여기서 한 걸음 더 나아간다면 최소 한 시간은 타이핑을 쳐 봐야 한다. 이런 훈련은 마치 운동 선수가 근육을 만들어 가는 과정과 매우 비슷하다. 뇌의 글쓰기 근육도 이렇게 만들어 가야 하는 것이다.

본격적인 책 쓰기 방법

이제 본격적인 책 쓰기 방법을 살펴보자.

처음엔 내가 어떤 책을 써야 할지 고민하는 시간이 상당히 길 것이다. 사실 이게 정상이다. 그런데 책은 내가 어떤 걸 써야지라는 생각으로 시작하기도 하지만 어느 날 갑자기 아이디어가 ‘뿜’하고 나타나기도 한다. 그런데 대부분 이런 아이디어들은 도로 뇌 속에 갇혀 영원히 밖으로 나오지 못하는 경우가 다반사다. 그래서 책을 쓰려고 마음을 먹었다면 그날부터 메모광이 되어야 한다.

메모를 하는 방법은 사람마다 각기 자기의 성격이나 습관에 따라 하면 된다. 하지만 그 타이밍은 지켜 주어야 한다. 일단 좋은 글감이 떠올랐다면 지체 없이 수첩이나 스마트폰의 메모장을 열어 기록해야 한다. 그리고 그 메모들을 째짤이 반복해서 읽어 주어야 한다. 이 과정은 사고를 확장하고 구체화하는 데 도움을 준다. 이런 메모와 확인 활동을 반복하다보면 점점 자신이 써야 할 책의 큰 그림이 그려진다.

책의 제목이나 글감이 확실히 정해졌다면 이제 그 표현하고자 하는 내용을 구조화해야 한다. 흔히 ‘시놉’ 혹은 ‘시놉시스’라고 하는데, 자신이 글을 어떻게 써나갈지에 대한 설계도와 같다. 설계도가 정확하면 이야기의 전개도 정확하게 진행이 되고, 설계도가 자세하면 이야기의 전개도 세밀하게 이루어진다. 또, 독창적인 설계도는 독창적인 글을 만들고, 화려한 설계도는 화려한 글을 만든다. 그만큼 시놉은 중요하다.

시놉에는 필수 요소가 있다. 이야기의 흐름인 기승전결이 잘 드러나야 한다. 그리고 캐릭터의 성격이 명확해야 한다. 또 독자의 흥미를 유발할 수 있는 반전이 하나 정도 있으면 좋다.

위의 조건에 맞는 시놉을 작성했다고 해서 무조건 그걸 토대로 글을 쓰는 건 아니다. 건축설계사가 설계를 했다고 모두 건물을 짓는 것은 아니듯 말이다. 일단 글을 쓰기 시작하면 생각보다 많은 시간과 에너지가 들어가기 때문에 이 작업에 들어가기 전에 최소한 세 가지 정도의 시놉은 짜보아야 한다. 그래서 그 중에 가장 쓰기에 좋고 매력 있는 시놉을 선택해서 글을 쓰는 게 좋다.

여기까지 왔으면 이제 다음은 여러분들의 역량이 알아서 글을 만들 것이다. 단 포기는 하지 않아야 한다. 그리고 글이 완성이 되면 퇴고를 한 후, 자기의 원고를 성실하게 읽어 줄 수 있는 사람을 한 명 이상 찾아 피드백을 받아야 한다. 작가들이 흔히 착각하기 쉬운 게 바로 내 머리 안에 있는 세계를 타인도 알고 있을 거라고 생각하는 것이다. 이런 부분들을 피드백을 통해서 교정하는 작업이 퇴고의 역할이다.

그리고 마지막으로 원고를 다 작성했다면 컴퓨터에 저장만 해두지 말고 꼭 투고를 해야 한다. 아무리 신에게 로또 1등에 당첨되게 기도를 하더라도 로또를 사지 않으면 안 되듯이 꼭 출판사에 투고를 해야 한다.

나를 작가의 길에 처음으로 불러준 편집장의 말 중에 공감이 되는 멋진 말 한마디로 책 쓰기에 대한 나의 글을 마무리하려고 한다.

“내 분야에서 전문성을 갖는 가장 멋진 방법은 바로 내가 쓴 책을 갖는 것이다.”



통계로 바라보는 세상이야기...

2020 Census 인구주택총조사 실시

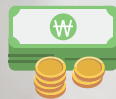
2020 Census
인구주택총조사 실시

인구주택총조사는 한 국가 또는 일정한 지역의 모든 사람, 가구, 거처와 관련된 인구·경제·사회학적 자료를 수집, 평가, 분석, 제공하는 전 과정으로 우리나라의 모든 사람과 주택의 규모 및 그 특징을 파악하기 위해 5년마다 실시하는 국가기본통계조사입니다. 인터넷(모바일)조사는 10월 15일부터 31일까지, 방문면접조사는 11월 1일부터 18일까지 실시하였습니다. 인구주택총조사는 역사적이고 범세계적 통계조사로 고대 바빌로니아(B.C. 3600년경) 시대에 기원을 두며, 근대적인 인구조사는 1790년 미국에서 최초로 실시하였습니다. UN 통계청에 따르면, 2010 라운드 센서스(2005~2014년)에서는 214개국에서 인구총조사를 실시하여 전세계 인구의 93%가 조사되었다고 합니다.

01

30년간 대한민국 경제는
어떻게 변했을까?

우리나라 경제는 1988년 서울 올림픽 개최 이후 크게 변했습니다. 1997년 말 IMF가 터지고 2008년 글로벌 금융 위기를 거치며 양적, 질적으로 큰 영향을 받았는데요, 'KOSTAT 통계플러스 불호'에 실린 서울시립대 경제학부 배영수 교수의 '주요 거시경제지표로 본 30년간 우리 경제의 변화'에 따르면, 한국의 경제 규모는 1988년 144조 원에서 2017년 1,730조 원으로 30년 사이에 12배 증가했지만, 1인당 명목 국민총소득(GNI)은 같은 기간 340만 원에서 3,364만 원으로 9.9배 증가하는 것에 그쳤습니다. 수출에서 수입을 뺀 '순수출'은 1988년부터 2007년까지 흑자였지만 2007년부터 2014년 8년간 적자였다가, 이후 회복



하여 2017년에는 65조 7천억 원의 흑자를 기록했습니다.

02

2019년 4인 가구
김장 비용은 얼마?

2019년 11월 기준, 4인 가구 김장 비용은 30만 원 내외로 예상됩니다. 이는 2018년 대비 약 10% 상승했습니다. 2019년 가을무 재배면적은 1,616만 6천 평(=5,344ha)으로 전년 대비 12.3% 감소했고, 가을배추 재배면적은 3,317만 8천 평(=10,968ha)으로 전년 대비 17.6% 줄었습니다. 이는 태풍, 잦은 강우, 전년 수확기 이후 가격 약세 영향으로 보입니다. 시도별 가을배추 재배면적은 전남(2,666ha), 경북(1,660ha), 충북(1,394ha), 경기(1,215ha), 강원(1,189ha) 순이며, 가을무 재배면적은 경기(1,224ha), 전북(1,006ha), 충남(806ha), 전남(796ha), 강원(463ha) 순으로 나타났습니다.

03

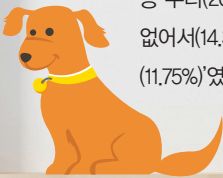
가격 폭락! 양파,
2019년 양파 생산량 때문?

올해 양파 가격이 크게 떨어졌습니다. 양파 농사가 잘 된 이유에는 날씨 영향이 큼니다. 지난겨울 따뜻한 기온과 함께 알이 굵어지는 시기에 적절한 강수량과 풍부한 일조량이 큰 영향을 끼쳤습니다. 통계청, '2019년 보리, 마늘, 양파 생산량 조사 결과'에 따르면 10a당 양파 생산량은 7,322kg으로 전년 5,756kg보다 27.2% 증가했습니다. 2000년 이후로 보면 2009년 7,412kg의 최고 생산량을 기록한 이후 두 번째로 높았고, 2010년부터 2018년까지는 5,000kg~6,000kg대의 생산량을 유지한 것과 비교해도 2019년이 특히 생산량이 높았습니다. 특히 2019년 양파 재배면적은 21,777ha로 전년 대비 17.6% 감소하였으나, 생산량은 159만 4,450톤으로 전년 대비 4.8% 증가했습니다.

04

반려동물등록제, 미등록 시 과태료 100만 원!

문화체육관광부의 2018년 '반려동물에 대한 인식 및 양육 현황 조사'에 의하면, 현재 반려동물을 기르고 있다고 응답한 비율이 27.9%이며, 응답자 중 81.3%(중복응답 결과)가 강아지를 가장 많이 기른다고 합니다. 반려동물 등록제는 3개월 이상의 반려견을 기르는 사람은 시, 군, 구청에 반려견을 등록하는 제도로, 반려견을 좀 더 쉽게 찾을 수 있도록 하고 유기견 발생을 줄이고자 하는 목적으로 시행하고 있는데요, 실제 반려견을 등록한 경우는 51.7%로 나타났는데, 등록하지 않는 이유 1위는 '필요성을 느끼지 못해서(36.8%)'였고, 2위는 '내장형 무선식별장치 삽입 후 부작용 우려(20.3%)', 3위는 '바쁘거나 시간이 없어서(14.3%)', 4위는 '절차가 번거로워서(11.75%)'였습니다.



05

인구 고령화, 노년의 경제·사회활동 참여는?

통계개발원의 'KOSTAT 통계플러스' 2019년 가을호에 실린 박시내 사무관의 '고령화와 노년의 경제·사회활동 참여'에 따르면 한국인이 '스스로 늙었다고 느끼는 평균 나이'는 60세로 이탈리아(70세)에 이어 두 번째로 높았습니다. 이는 세계 41개국 평균인 55세보다 5세 높은 나이인데요, 우리 사회가 고령사회로 접어들며, 고령기의 역할을 새롭게 규정하려는 움직임도 나타나고 있습니다. 2017년 기준 OECD 국가 고령층(65세 이상)의 고용률을 살펴보면, 인도네시아가 41.1%로 가장 높으며 아이슬란드(38.2%), 한국(30.6%), 콜롬비아(29.0%), 멕시코(26.6%) 순으로 나타났습니다. 특히 OECD 국가들의 고령층 평균 고용률은 14.4%로 대한민국은 이보다 약 2배 높은 수준입니다.

06

통계개발원이 알려주는 한국의 'SDGs'

세계적으로 주목받고 있는 SDGs(Sustainable Development Goals), 즉 지속가능발전 목표에 대해 통계개발원 연구기획실 박영실 사무관의 「한국의 SDGs 이행현황 분석」을 살펴보면, SDGs는 인류가 공동으로 달성해야 할 발전 목표로 17개 목표와 169개의 세부 목표가 있습니다. 이 중 SDG4(교육) 부문의 읽기 및 수학 영역에서 최소성취수준을 달성한 한국 청소년의 비율이 각각 86.3%와 84.5%로 나타났고, SDG7(에너지) 부문에서 2016년 기준 전 세계 재생에너지 비율은 17.5%로 나타났으며 SDG11(도시) 부문에서 대기오염은 중요한데요, 한국의 초미세먼지 농도는 $24.6\mu\text{g}/\text{m}^3$ 으로 OECD 회원국 중 한국보다 농도가 높은 국가는 터키($42.0\mu\text{g}/\text{m}^3$) 뿐이었습니다.



07

10대 20대도 걸린다는 '디지털 치매'

디지털 기기에 지나치게 의존해 기억력과 계산 능력이 크게 떨어지는 상태를 '디지털 치매'라고 합니다. 2004년에 국립국어원에 신조어로 등재된 '디지털 치매'는 등장한 지 오래였지만, 2019년 트렌드몬이터의 '디지털 치매 관련 인식 조사'에 따르면 디지털 치매를 들어본 적은 있지만, 내용에 대해 잘 모른다고 응답한 비율이 37.9%, 디지털 치매를 전혀 모른다고 응답한 사람이 27.5%를 기록했습니다. '외우는 전화번호가 몇 개 되지 않는다' '애창곡 가사를 보지 않으면 노래를 부를 수 없다' 등 10개의 디지털 치매 자가 체크리스트 상 4개 이상 항목에 공감해 디지털 치매 해당자에 속한 사람은 36.2%, 3개 항목에 공감해 디지털 치매 위험군에 속한 사람은 22.9%를 기록했습니다.



08

통계로 들여다본 우리의 언어예절

2015년에 국립국어원에서 실시한 '국민의 언어의식조사'를 보면, 국민이 가장 관심을 보이는 국어 분야로 '언어예절'이 70.2%로 1위를 차지했습니다. 언어예절의 뒤를 이어서는 말하는 태도나 방법과 같은 '말하기' 분야가 69.0%를 기록했습니다. 동방예의지국이라고 불려온 나라답게, 1위와 2위 모두 타인과 대화할 때의 예절과 태도와 관련이 있었습니다. 일상 속에서 글을 쓸 때나 말할 때 적잖이 신경 쓰이는 분야인 '맞춤법과 발음'도 55.2%로 높은 비율을 차지했습니다. "손님, 주문하신 커피 나오셨습니다." 사람이 아닌 사물을 높여 말하는 현상인 '사물 존칭'입니다. 이는 언어예절을 과하게 지키다 발생한 현상인데요, 사물 존칭에 대한 인식은 부정적 의견이 훨씬 높게 나타났습니다.

09

클라우드 펀딩, 꼼꼼하게 따져보자!

클라우드 펀딩은 대중을 뜻하는 크라우드(crowd)와 자금 조달을 뜻하는 펀딩(funding)이 합쳐진 용어로, 자금을 필요로 하는 수요자가 온라인 플랫폼 등을 통해 불특정 다수 대중에게 자금을 모으는 방식입니다. 트렌드몬итор의 '2019 클라우드 펀딩 관련 인식 조사'에 따르면 클라우드 펀딩이라는 용어를 들어본 사람은 78.2%로, 그중 직접 참여해 본 경험이 있는 사람은 14.5%로 나타났습니다. 클라우드 펀딩에 참여할 때에 주의할 점도 있습니다. 주요 투자대상이 벤처, 스타트업이기 때문에 투자위험이 크고, 투자금을 회수하지 못할 가능성도 있습니다. 따라서 투자 대상 기업에 대해 꼼꼼히 살펴보고, 사업 진행 상황에 대해서도 자주 확인하는 것이 필요합니다.



10

직장 내 괴롭힘 금지법, 시행 세 달간의 기록!

지난 8월, 고용노동부는 직장 내 괴롭힘 금지법이 시행된 이후 1개월간 접수된 진정을 분석한 결과를 공개하였는데, 접수된 진정은 총 379건(19.8.16 기준)으로 하루 평균 16.5건에 달하는 수치라고 합니다. 먼저 지역별로 보면, 서울 119건, 경기 96건으로 전체 진정의 56.7%를 차지하였으며, 유형별로는 폭언에 관한 진정이 152건(40.1%)으로 가장 높았고, 그다음으로는 부당업무지시(28.2%), 험담·따돌림(11.9%) 등의 순서로 나타났고, 규모별로는 50인 미만 사업장 소속 근로자에게서 접수된 진정이 159건(42.0%)으로 압도적으로 높은 비율을 차지했습니다.



11

국가기관 최애하는 20대의 사회의식은?

2006년 7월, 국가통계 전문 연구기관으로 탄생한 통계개발원의 박나라 주무관이 연구한 '20대의 사회의식과 여가행태 변화'에 따르면, 2017년 단체활동 참여 비율에서 20대는 취미여가활동 참여 54.5%, 학술 단체 참여 13.0%로 높은 수준을 보였고, 기부 참여율은 2017년 20.7%로 30대(31.2%), 40대(36.5%)에 비해 낮았습니다. 최근 확산된 금수저, 흙수저 등 수저계급론에 대해 20대는 본인의 사회·경제적 지위가 '중'이라고 응답한 사람은 2017년 62.1%였습니다. 2017년 20대가 직업을 선택할 때 가장 중요시하는 요인은 수입이 33.1%로 1위를 차지했고, 선호하는 직장으로는 국가기관이 24.9%로 가장 높았으며, 공기업 22.8%, 대기업 13.7%, 자영업 13.5% 순으로 나타났습니다.

12

독도명예주민증 발급 5만 명 돌파

1900년 10월 25일, 고종황제가 대한제국칙령으로 처음 독도를 울릉도의 부속 섬으로 명시한 날이라, 10월 25일이 독도의 날이 되었다고 합니다. 또한 이 행정조치로 인해 일본이 '1905년 주인 없는 땅인 독도를 취득했다'는 주장에 대해 '당시 독도는 주인 없는 땅이 아니었다'고 반박할 수 있는 중요한 근거가 되었습니다. 1965년 우리나라 최초의 독도 주민인 최종덕 씨를 시작으로, 가족관계등록부상으로 집계된 독도 인구수는 2003년 842명에서 15년 후 2018년에는 3,437명으로 약 4배나 증가했습니다. 지난 2013년에는 25



만 명이 찾아 최대 방문객 수를 기록했고, 작년도에도 이에 못지않은 22만 명이 방문했으며, 작년 12월 기준으로 방문객 수 누계가 220만 명을 돌파했습니다.

13

통계와 함께 떠난 박씨네 등산 VLOG

통계청 대학생 기자단의 박희인 기자가 가족들과 함께 지리산을 찾았다고 합니다. 여성가족부의 2015년 가족실태조사에 의하면, 휴일 가족 여가 활동 내용(1순위)으로 등산하는 비율이 4.2%였습니다. 그 이유로는 자연을 보며 스트레스를 풀 수 있으며 동시에 운동을 해 건강을 챙길 수 있다는 점을 꼽았습니다. 문화체육관광부 2018년 국민생활체육조사를 보면, 체육활동 참여 이유(1순위)로 42.9%가 건강 유지 및 증진(1위)을 위해 체육활동에 참여한다고 응답했습니다. 특히 가족 간의 관계를 친밀하게 하는 데도 도움이 된다고 합니다. 휴일을 맞아 가족과 함께 등산하며 즐거운 시간을 보낸 통계청 대학생 기자단. 부모님께 인터뷰도 진행했는데요, 어떤 이야기를 주고받았을까요?

14

'텍본 공유'... 인터넷상에 퍼지는 불법복제물

'텍본'이란 컴퓨터에 내장된 '메모장'이라는 프로그램으로 저작권 작품을 직접 작성하여 생성하는 '불법복제물'이며 이는 웹소설이나 문학 작품 등 '출판' 콘텐츠에 있어서 가장 큰 피해를 주는 복제물입니다. 한국저작권보호원의 2019 저작권 보호 연차보고서에 따르면, 2018년 한 해 동안 불법복제물 이용경험 인구(만 13세~69세)는 약 1,773만 명으로 전체 인구의 43.4%가 불법복제물을 이용한 것으로 나타났습니다. 불법복제물의 이용은 2013년 이후 계속해서 증가하다가 2015년 소폭 감소 후 2016년 다시 증가했습니다. 2018년 기준, 합법저작물 시장을 침해하는 콘텐츠 중 음악



이 약 63% 가장 많은 비중을 차지하고 있으며 그 수가 12억 6천만 개에 달합니다.

15

해외여행 감염병 예방 준비도 통계로!

보건복지부 질병관리본부의 2018년 감염병 감시연보에 의하면, 법정감염병 국외 유입 환자가 2014년 400명, 2015년 491명, 2016년 541명, 2017년 531명 그리고 2018년 597명으로 증가하는 추세를 보입니다. 2018년에 국외에서 유입된 감염병은 Dengue열(26.6%)이 가장 많았고 세균성이질(24.3%), 장티푸스(15.4%), 말라리아(12.6%), C형 간염(4.7%)이 그 뒤를 이었습니다. 2018년 우리나라 Dengue열 환자는 159명으로 전년 대비 7% 감소하였으며, 세균성이질에 걸린 환자는 2018년 총 191명으로 2017년에 비해 70.5% 늘었고, 장티푸스 환자 또한 국외유입 환자 증가로 2018년 전년 대비 66.4%가 증가하였으며, 말라리아 환자는 2018년 576명으로 전년 대비 11.8% 증가하였습니다.



16

통계청, 정부부처, 지방자치단체, 연구기관 등
모든 기관의 마이크로데이터를 한 곳으로...

보다 심도 있고 다양한 분석을 원한다면
지금 바로 **MDIS**를 클릭해보세요

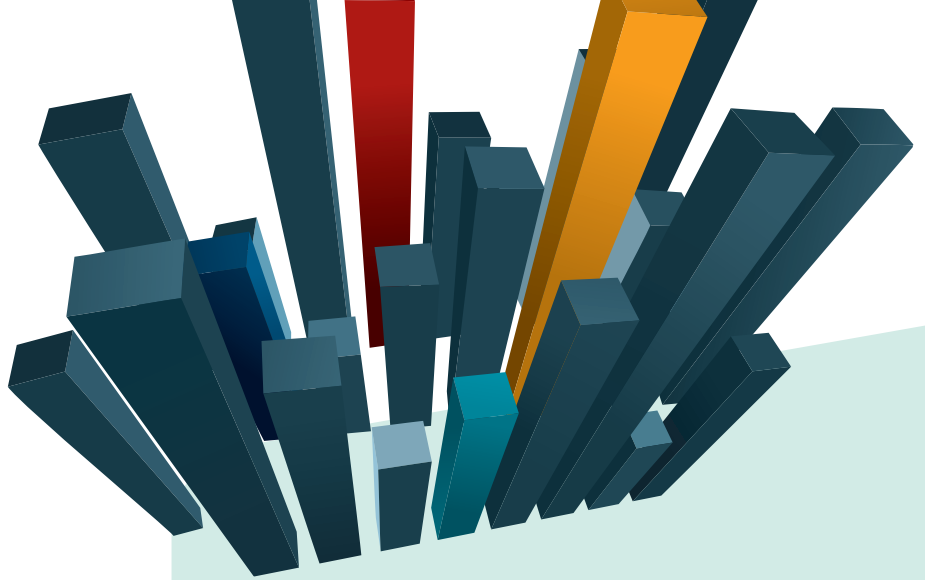
■ 서비스 소개

가. 서비스명 : 마이크로데이터통합서비스(MDIS, mdis.kostat.go.kr)

나. 제공통계수 : 총 202종(통계청 45종 및 통계작성기관 157종)

다. 제공형태 : 마이크로데이터(통계에 따라 사람, 사업체, 가구 기반 자료)

기준		주요 통계
통계청	인구 · 가구	경제활동인구조사, 가계동향조사, 국내인구이동통계, 사망원인통계, 가계금융복지조사, 지역별고용조사, 인구주택총조사, 인구동향조사, 생활시간조사, 사회조사 외 7종
	사업체 · 농어가	전국사업체조사, 광업 · 제조업조사, 농가경제조사, 기업활동조사, 농업총조사, 농산물생산비조사, 경제총조사, 어가경제조사, 운수업조사 외 11종
	행정통계 및 기타	귀농귀촌인통계, 영리법인기업체행정통계, 신혼부부통계, 주택소유통계, 퇴직연금통계, 일자리행정통계, 국부통계조사
통계작성기관		전국다문화가족실태조사, 가족실태조사, 자동차주행거리통계, 직종별사업체노동력조사, 보육실태조사, 지상기상통계, 국민여가활동조사, 외래관광객실태조사, 한부모가족실태조사, 청소년종합실태조사 외 147종



■ 서비스 내용

가. 구분: 자료의 민감성 정도에 따라
공공용, 인가용, 특수목적용으로 구분 운영

나. 수수료

- 무료: 공공용 자료
- 인가용/특수목적용: 선택제 수수료 부과

다. 서비스 방법

- 추출·다운로드: MDIS 포털에서
직접 무료 다운로드
- 원격접근서비스: 승인 후 이용자가
집·사무실 등에서 통계청 서버 접속 후 활용
- 이용센터: 승인 후 지정된 장소를 방문·활용

■ 문의

- 연락처: 재단법인 한국통계진흥원
- 전화: (02) 512-0167 FAX: (02) 515-0240
- 주소: (우) 06097 서울특별시 강남구 선릉로 612, 6층
- E-mail: MDIS@stat.or.kr

공공자료의 개방 및 공유 확대

MDIS
[mdis.kostat.go.kr]

통계청 46종 및 통계작성기관 124종의
통계자료로 제공확대

통계청에서 국가통계를 활용하세요!

통계청은 통계개발·활용에 필요한 모든 정보와 도움을 제공합니다.

다양한 국가통계정보 제공사이트를 활용하세요.

원하는 자료를 직접 분석 및 요청

MDIS
[mdis.kostat.go.kr]

온라인으로 추출/다운로드 선택 시
공공용 마이크로데이터를 무료로
분석 활용 가능



국가통계 쉽게 찾기

KOSIS
[kosis.kr]

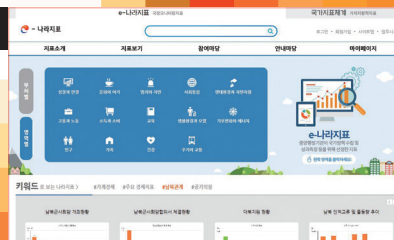
국내, 국제, 북한의 주요 통계를
한 곳에 모아 알기 쉽게 분류해 제공



국가 발전 상황을 한눈에

국가지표체계
[www.index.go.kr]

국민의 관심이 크고 정책 수립에
활용 가능한 지표



지도 위 통계정보 살펴보기

SGIS
[sgis.kostat.go.kr]

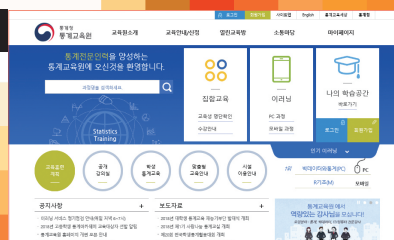
인구, 가구, 주택, 사업체 통계 등 각종
통계를 지도(GIS) 위에서 한눈에 파악



국내 유일의 국가통계교육 전문기관

통계교육원
[sti.kostat.go.kr]

통계작성 및 활용 전문통계과정,
기관맞춤형과정, e-러닝 과정



통계청
통계교육원