

데이터 기반

의사결정 시대의 통계

SPRI Spring 컨퍼런스
2019. 4. 12

정보통신정책연구원(KISDI)
데이터사이언스그룹장 정용찬

- 이 자료는 '증거기반 의사 결정 강화를 위한 국가통계 거버넌스 발전방안(국가통계발전포럼, 2018. 8), '4차 산업혁명 시대의 데이터 경제 활성화 전략'(KISDI Premium Report 17-04, 2017. 6. 14)과 '4차 산업혁명 시대의 데이터 거버넌스 개선 방향'(KISDI Premium Report 18-05, 2018. 8. 7)을 기초로 작성함

차 례

1. 데이터 시대의 의사결정
2. 주요국의 데이터 거버넌스 전략
3. 데이터 시대의 통계

Timeline of statistics

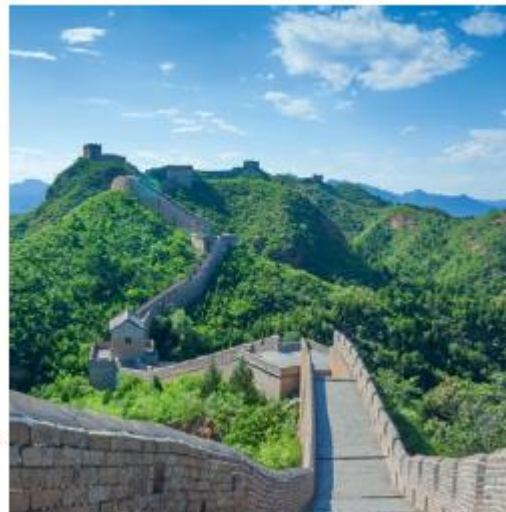
AD 7 Census by Quirinus, governor of the Roman province of Judea, is mentioned in Luke's Gospel as causing Joseph and Mary to travel to Bethlehem to be taxed.



Roman Gudyma/iStock/Thinkstock



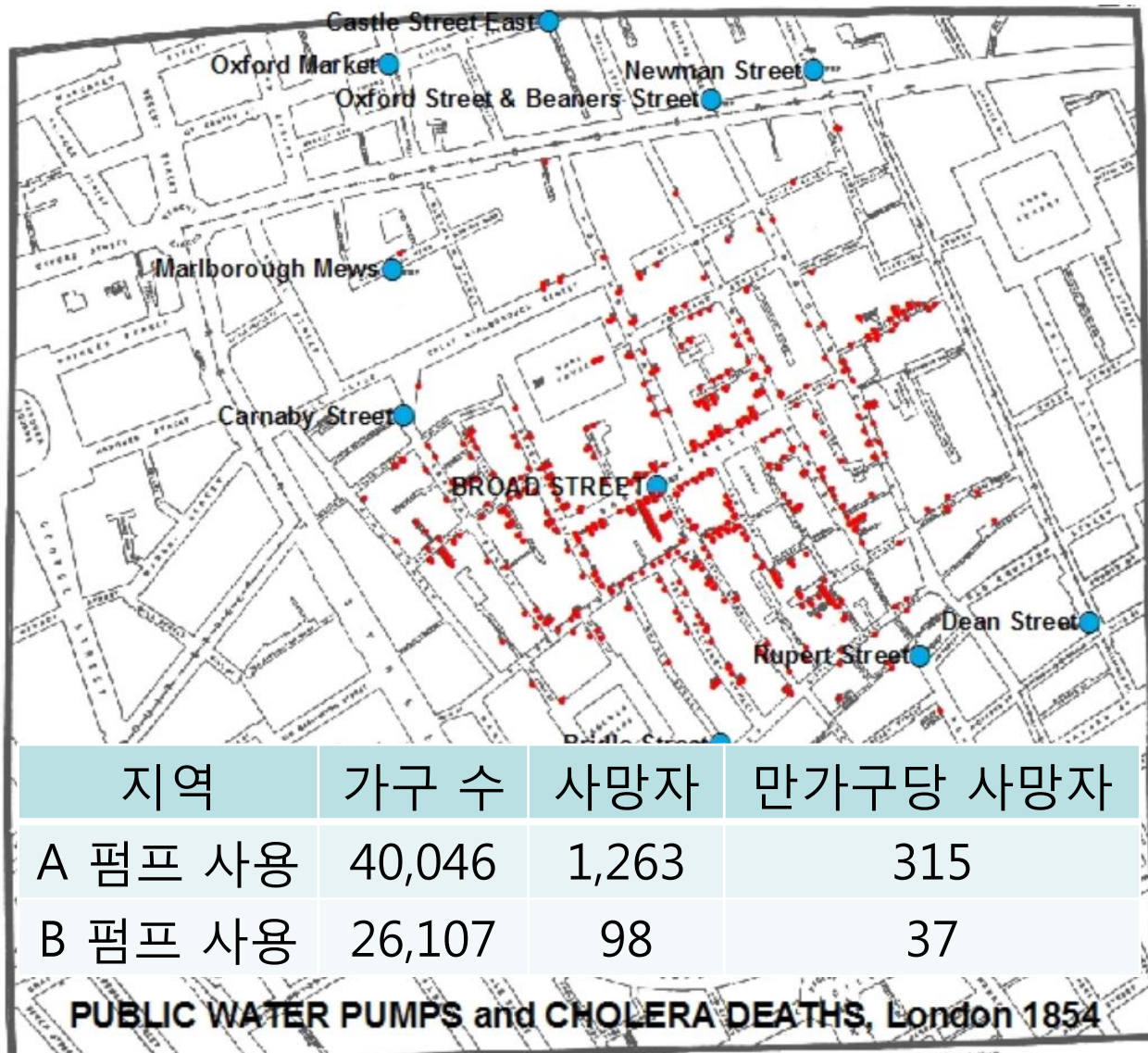
AD 2 Chinese census under the Han dynasty finds 57.67 million people in 12.36 million households – the first census from which data survives, and still considered by scholars to have been accurate.



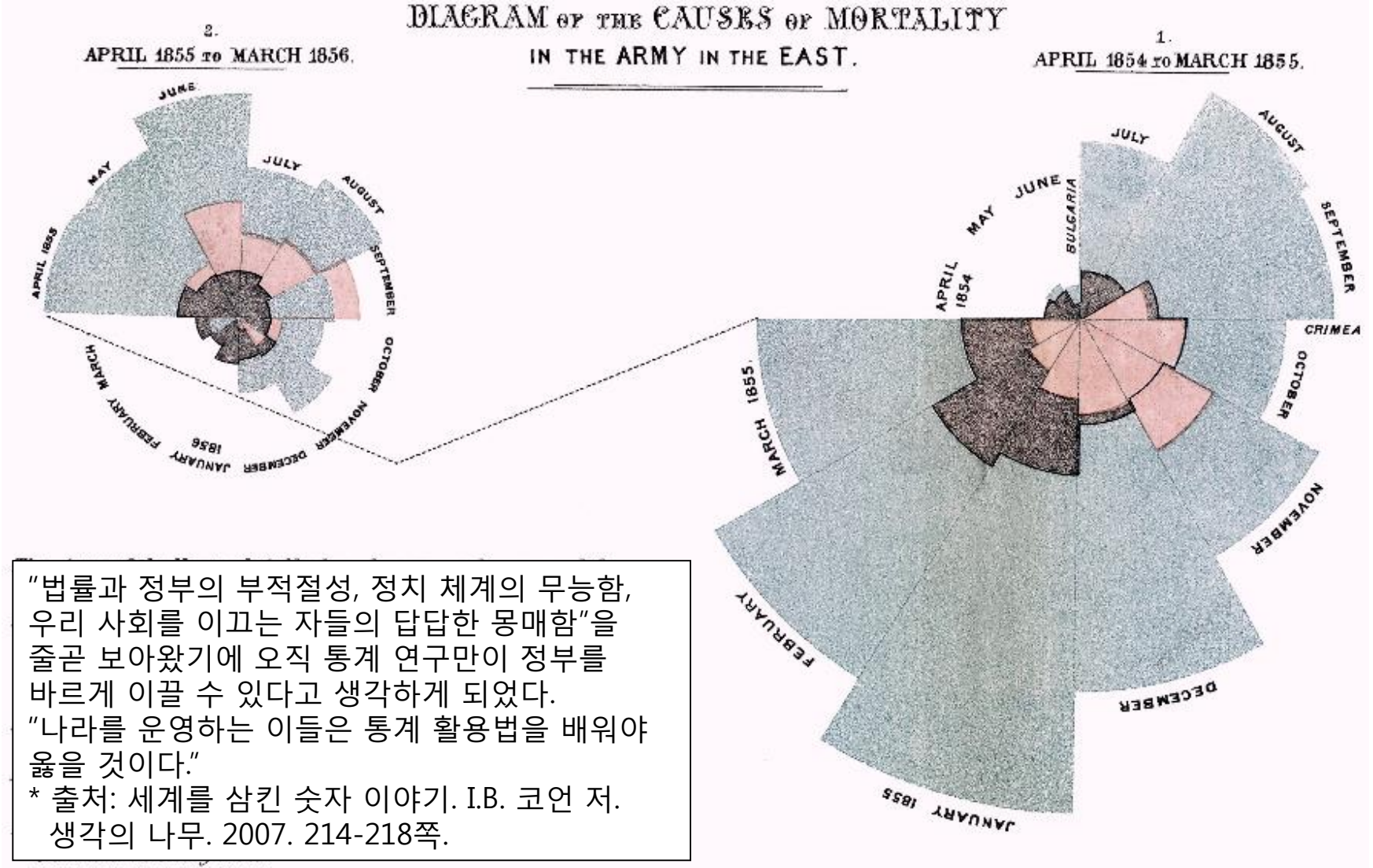
iStock/Thinkstock



콜레라를 막은 Snow 박사의 데이터

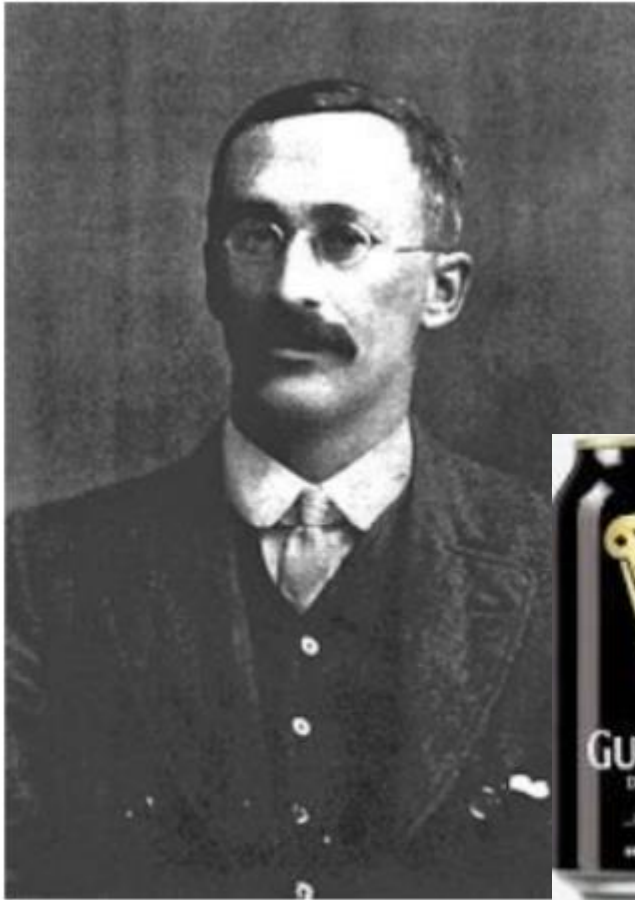


나이팅게일의 Rose Diagram



<https://commons.wikimedia.org/wiki/File:Nightingale-mortality.jpg>

기네스맥주와 t-distribution



William Sealy Gosset (1876–1937)

Plaque in the Guinness Storehouse
Commemorating Gosset.

Famous as a statistician, best known by his pen name *Student* and for his work on Student's t-distribution.

Hired by Guinness to find a cheap way of monitoring stout quality

<https://bioinformaticsandme.tistory.com/80>
https://en.wikipedia.org/wiki/William_Sealy_Gosset

● 세종12년(1430) 3월~8월 여론조사

- 토지 비옥도에 따라 6등급, 풍년과 흉년을 기준으로 9등급
- 세금 제도 개선 여론조사: "정부, 육조와 각 도의 감사, 수령 및 품관으로부터 여염의 세민에 이르기까지 새로운 조세제도에 대하여 모두 가부를 물어서 아뢰게 하라." 찬성 98,657(57%) 반대 74,149
- 세종 19년 압도적으로 찬성한 경상, 전라 시범 시행, 세종 26년 전국 시행



- 재판에서 사실의 인정은 증거능력이 있는 증거에 의하여 행하여야 한다(형사소송법 307조)



<https://movie.naver.com/movie/bi/mi/basic.nhn?code=35901>

<https://movie.naver.com/movie/bi/mi/scriptAndRelate.nhn?code=101901#scriptMovie>

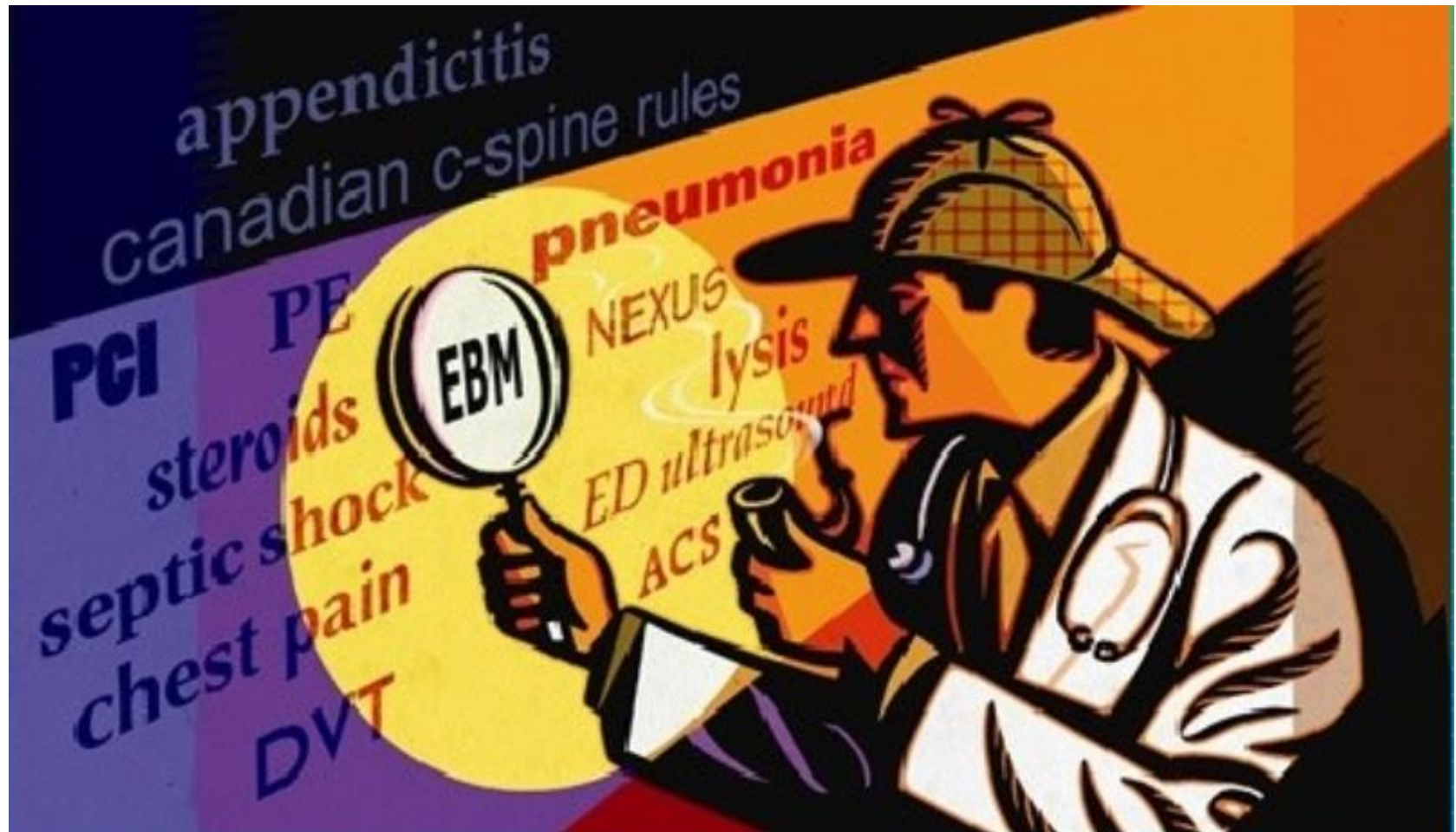


● FBI CODIS: 등록된 범죄자 DNA로 용의자 색출

- 물잔, 껌, 봉투, 혈흔, 권총 방아쇠, 모자
- DNA Identification Act 의회 통과 (1994)
- National DNA Index System(NDIS) 구축 (1998)
- DNA Fingerprint Act (2005)
: 50개 주 가입, 범죄자(1256만)와 체포자(250만) DNA 정보 보유(2016년 9월)
- 실종자 확인 분야로 확대 예정

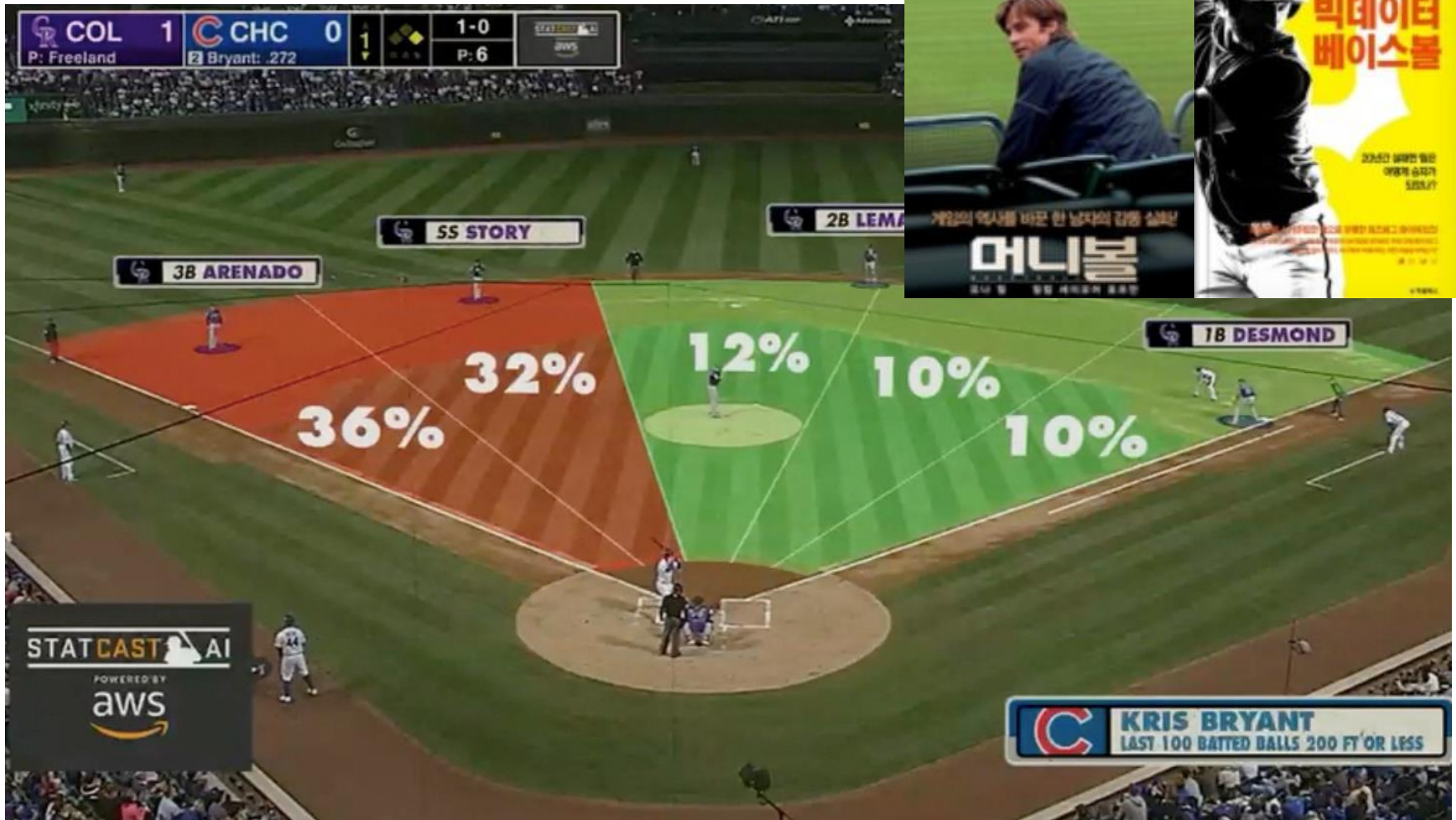


- approach to medical practice intended to optimize decision-making by emphasizing the use of evidence from well-designed and well-conducted research



야구통계 Sabermetrics

● MLB STATCAST: 7.3 tera per game



● 데이터 기반 인사관리

- 인사관리부서: People Analytics Team

● 승진 확률: 90% 정확도

$$\text{Odds} = e^{-22.216 + (5.227 \times \text{AvgPerf}) + (2.732 \times \text{MgrRecommended}) + (.971 \times \text{SelfRecommended})^*}$$

$$\text{Probability (\%)} = \frac{\text{Odds}}{1 + \text{Odds}}$$

Source: <http://www.businessinsider.com/google-promotion-equation-2014-11>

● Project M&M

- 공짜 간식 제공: 초콜릿 등 고칼로리 간식과 탄산음료 과다 섭취
- 초콜릿은 불투명 용기에, 건강 스낵은 투명 용기에. 탄산음료는 냉장고 아래 쪽에, 생수는 눈높이 위치에: 설탕 첨가 음료 7% 감소, 생수 소비 47% 증가
- 뉴욕 사무실 7주 동안 총 310만 칼로리 감소(초콜릿 자판기 9대 분량)

Source: Cecilia Kang, 'Google crunches data on munching in office', Sep. 1, 2013, Washington Post.



모든 데이터를
사람들에게



포렌식 아키텍처 Forensic Architecture
수퍼플렉스 SUPERFLEX
자크 블라스 Zach Blas
조오 페이 Coo Fui
크리스 셴 Chris Shen
하임 반 덴 도르펠 Harm van den Dorpel
레이첼 아라 Rachel Ara
사이먼 웨니 Simon Denny
김설비 Syibee Kim
김우현 Woonghyun Kim

2019.3.23 - 7.28

VERTIGINOUS DATA
국립현대미술관
서울
3, 4전시실

NATIONAL MUSEUM OF
MODERN AND
CONTEMPORARY ART
SEOUL
GALLERY 3, 4

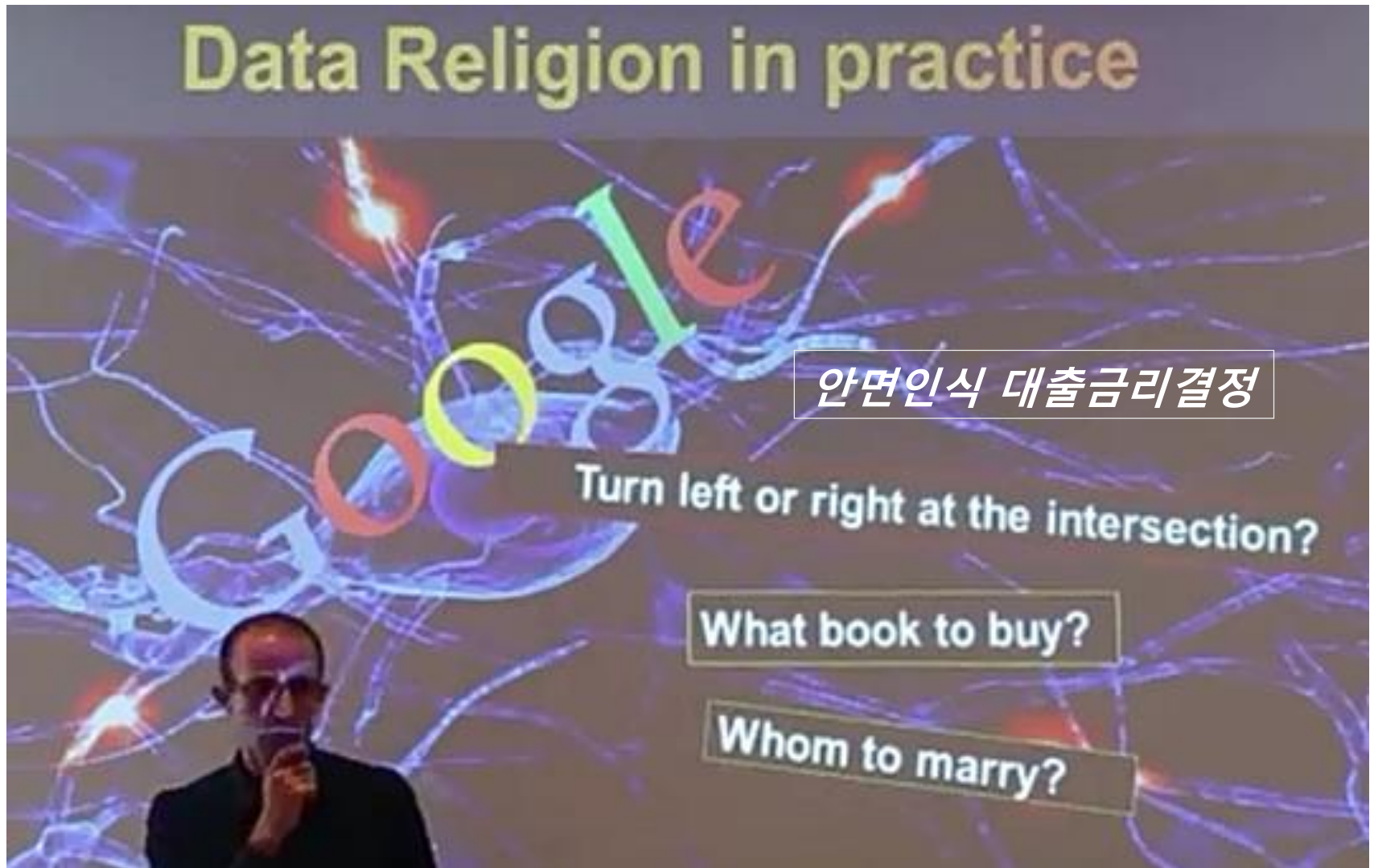
MMCA SEOUL

붉은한 데이터

데이터를 가공, 소유,
유통하는 주체는
누구이며, 어떠한
방식으로 그들이 가진
정보를 권력화 하는
것인가. 데이터를
둘러싼 맹목적인 믿음,
또는 그 근거 없는
불신과 위기감은
개인의 삶에 어떤
영향을 미치는가.
공공의 선에
기여하도록 데이터를
활용하는 것은 과연
가능한 것인가. 이번
전시는 이러한
질문들을 바탕으로
디지털 기술을
탐구하고 미적 특징을
발견하는 예술가들의
작품을 선보인다.

- 데이터: 화폐나 부동산, 지적재산권과 같은 새로운 '자산'





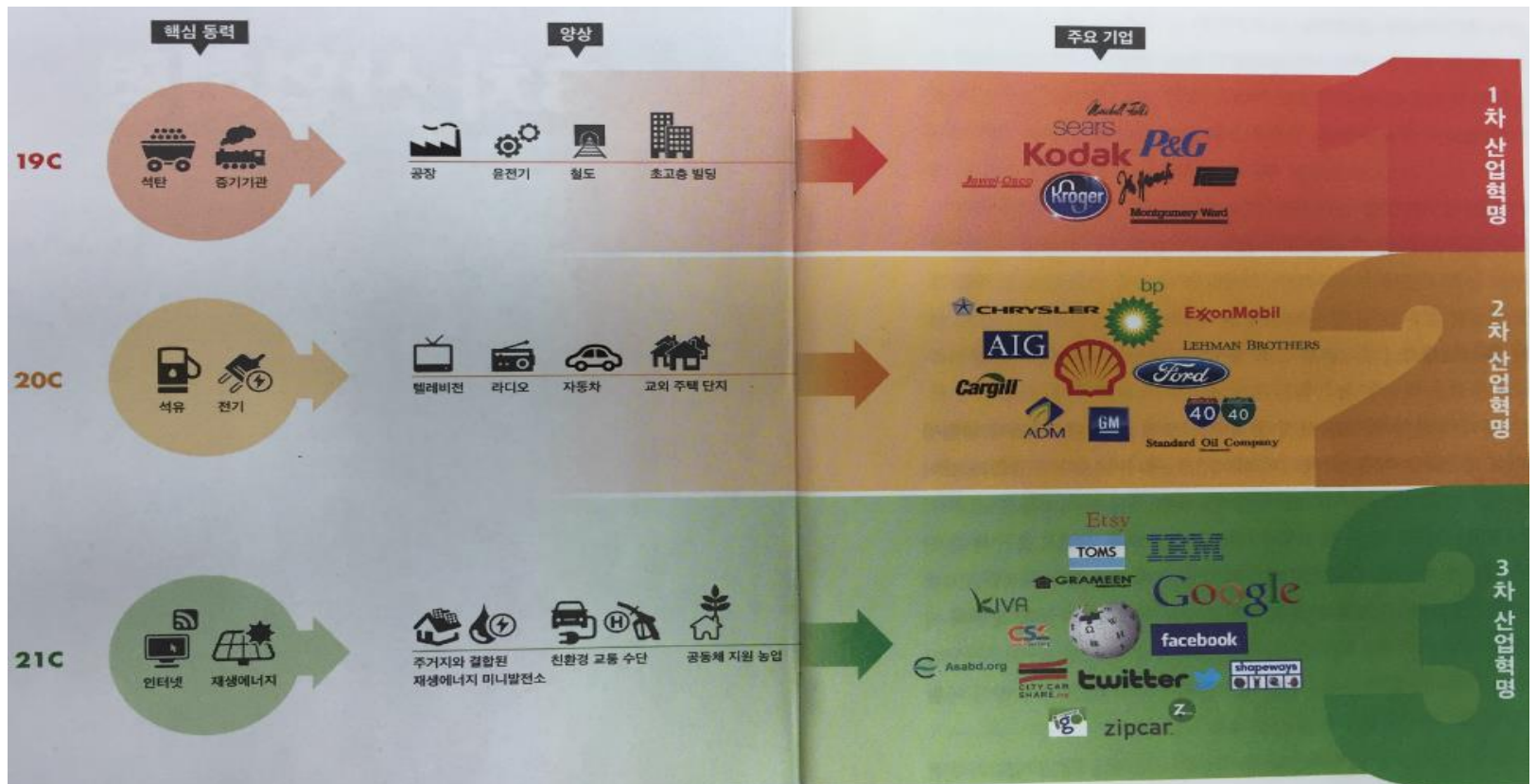
산업혁명과 시가총액

● 2019: 아마존, MS, 알파벳(구글), 애플, 페이스북, 알리바바, 텐센트

- 2007: 1.페트로차이나, 2.엑슨, 3.GE, 4.차이나모바일, 6.MS, 7.가즈프롬, 9.AT&T





- 1997: 1.GE, 2.로열더치셸, 3.MS, 4. 엑슨모빌

※ 한국 재계 순위(2018): 1.삼성, 2.현대, 3.SK, ... , 39.카카오, 49.네이버, 52.넥슨



데이터 경제(Data Economy) 수직계열화 용이 1. 데이터 시대

- 데이터 관련 협업 과정에서 데이터 생산, 인프라 제공, 연구조사 등을 담당하는 구성원으로 이루어진 생태계(European Commission, 2017)

데이터 표현자 (Data Presenters)	사용자 인터페이스 (User Interface)		사용자 경험 (User Experience)
	조사와 발견 (Investigation and Discovery)		사용자 관여 (User Engagement)
통찰력 제공자 (Insight Providers)	통계방법론 (Statistical & computational methods)	분석개발환경 (Development environment for analytics)	알고리즘/로직/규칙 (Algorithms/ logic/rules)
	시맨틱 모델 (Semantic model)	분석 라이브러리 (Analytics library)	머신러닝 (Machine Learning)
플랫폼 소유자 (Platform Owners)	개발 환경 (Development environment)		앱용 클라우드/호스트 (Cloud/host for apps)
	연계형 API (APIs for connectivity)		장치발견절차 (Device discovery)
데이터 수집자/ 데이터 관리자 (Data Aggregators Data Custodians)	데이터 표준화 (Data normalization for common transmission)		이종 데이터수집 (Heterogeneous data collection from disparate devices)
데이터 생산자 (Data Producers)	데이터 접근 (Data access)	데이터 통제 (Data control)	데이터 수집 (Data collection)
조력자(Enablers)			
   			
네트워크		하이브리드 클라우드	

데이터 기업의 시장 지배력

1. 데이터 시대

구분	게임 체인저	파급 효과
구글	AdSense 유튜브	광고 시장 신문, 방송
페이스북, 트위터	SNS	통신 시장
넷플릭스	추천서비스	비디어 대여 시장 방송콘텐츠 시장
아마존	AWS(클라우드)	컴퓨터 장비 방송콘텐츠 시장
5G 기반 데이터 기업	?	??

출처: 정용찬(2018). 5G 시대 글로벌 방송통신 환경변화, 율촌 방송통신산업 발전방안 세미나. '18. 6. 28.

● 데이터 경제의 부상과 증거기반 정책(Evidence-based policymaking)

- 데이터가 화폐나 부동산, 지재권에 견줄 만한 새로운 '자산'으로 주목
- 주요국은 '열린 정부(Open Government)'와 '오픈 데이터(Open Data)'를 표방, 데이터에 기반한 행정은 물론 데이터 공유와 활용에 사활
- * 정부 정책에서 데이터가 중요한 증거기반 환경에서 **통계의 신뢰성은 중요**

● 통계 생산 환경 변화

- 조사 거부, 불응 등 응답률 저하 심각
- **가계동향조사** '07 17%에서 '14 22.5%, 인구총조사 '10 0.23%에서 '15 0.88%
- 행정자료(인구주택총조사), 신용카드데이터(물가지수) 활용 시도
- 네덜란드 통계청의 도로 센서 빅데이터 활용 교통통계, 이탈리아의 국가통계청 자료와 이동통신 자료 혼합한 소지역 빈곤 추정치, 캐나다 통계청은 설문조사 대신 위성사진자료를 농업 통계 입력값으로 사용. 호주통계국은 위성표면 반사율 자료를 이용해 작물 유형을 분류하고 생산량을 추정(National Academy of Sciences, 2017)



영국

정부혁신전략보고서(Government Transformation Strategy 2017-2020, 2017)

- 세계 최고 수준의 디지털 서비스 제공, 정부 운영 방식을 효율적으로 혁신
- 의사결정 시 데이터 활용 위한 정부의 데이터 사이언스/분석 능력 강화

디지털경제법(Digital Economy Act, 2017)



미국

증거기반정책수립 보고서(The Promise of Evidence-based policymaking, 2017)

- 정책결정을 뒷받침 하는 증거생성 능력 강화 포괄적 전략 제시
- 국립데이터서비스청(National Secure Data Service, NSDS) 신설

연방통계 혁신 보고서(Innovations in Federal Statistics, 2017)



일본

미래투자전략 2018-Society 5.0 데이터추동 사회로의 혁신

- "데이터 패권주의" 로 경제사회 시스템의 건전한 발전 우려
- 정부 보유 데이터를 개방, 데이터를 활용한 혁신과 뉴비즈니스 창출 촉진

통계개혁추진회의, 2017

● Statistics and Registration Service Act 2007

- 미래지향적 **비부처 독립 통계위원회 제안**
- * 공공재(Public Good)인 공식통계(Official Statistics) 생산
- * 공식통계의 품질 보장과 포괄 범위 확장

● UK Statistics Authority, 2008

- Statistics and Registration Service Act 2007에 근거 설립
- 의회 지원 독립 기구인 위원회의 주요 기능
- * **모든** 공식통계에 대한 모니터링과 보고 체계, 독립적인 평가
- * 통계청(Office for National Statistics) 관리 감독
- **부처와 위원회 간의 통계적 목적의 데이터 공유 원칙 제시**

● Code of Practice for Official Statistics, UKSA(2009)

- 제6원칙: **비용 관점의 균형 감각(Proportionate burden)**
- * 데이터 생산자는 생산 부담이 통계 이용에 따른 효용을 넘어서는지 판단
- 실행전략 1. 통계 생산 비용 추정을 매년 보고하고 비용 절감 방법 개발
- 2. 공권력 의지보다는 협조에 의한 통계 조사 추구
- 3. **행정자료의 통계적 목적을 염두에 둔 적극적 행정 체계 설계**

● 디지털경제법(Digital Economy Act 2017)

- 전자 통신(electronic communications) 인프라 및 서비스에 관한 규정 마련
- 통계 작성 목적의 **민간 자료 이용 요구 조항**
- 연구 목적의 자료 공유, 정부/공공 기관 정보 공개, **통계 작성을 위한 정보 공유 조항이 포함**
- * 법 제정 이전에는 자료 공유 필요 시 'Statistics and Registration Service Act 2007'을 매번 개정

School census, National student database	Department for Education	24 개월
Student demographic information	Higher Education Statistics Agency	22
Pupil level school census for Wales	Welsh Government	18
VAT Information	HM Revenue and Customs	20
Customer Information System data	HM Revenue and Customs, Dept for Work and Pensions	23

UKSA(2016). Delivering better statistics for better decisions, Why we need new legislation for better access to data.

● 정부 혁신 전략(Government Transformation Strategy) 2017-2020

- 정부 혁신 전략의 목표

- * 세계 최고 수준의 디지털 서비스 제공, 정부 운영 방식을 효율적으로 혁신
- * 정책 과정에 시민과 정부가 함께 참여하는 적절한 기술 개발과 문화 조성
- * 공무원의 효과적인 업무 수행 위해 업무 환경과 프로세스 개선
- * 투명성 확보와 **정부와 민간의 혁신이 가능하도록 데이터의 활용 강화**
- * 공유플랫폼을 활용, 디지털 전환 가속화

- 데이터 활용 전략

- * API를 활용, 정부 데이터의 대내외 공개
 - * 효과적인 데이터 활용을 위해 **정부 조직 간의 장벽 제거**
 - * 데이터 활용을 선도하기 위한 정부 내 **Chief Data Officer 임명**
 - * 범정부 데이터 활용 위한 데이터 자문위원회(Data Advisory Board) 신설
 - * 의사결정 시 데이터 활용 위한 **정부의 데이터 사이언스/분석 능력 강화**
 - * 데이터 공유의 윤리 준수 위한 데이터의 안전하고 적절한 관리
- ※ 영국은 2016년 UN 전자정부 조사에서 1위(호주 2위, 한국 3위)



● 연방통계 혁신(Innovations in Federal Statistics: Combining Data Sources While Protecting Privacy), 2017

- 연방통계 생산 위한 표본 조사는 비용 상승, 응답률 하락으로 위기
- 연방기관 보유 **행정 기록(administrative records) 활용** 필요
 - * 행정기록 활용 위한 품질과 유용성 진단 필요
 - * 기관간의 공유 금지 등 법적, 행정적 규제 개선 필요
 - * 조사 자료와 행정 자료의 결합을 통해 데이터의 가치를 높이는 시도 필요
- 빅데이터 환경에서 **민간 자료를 활용한 연방통계** 작성 필요
 - * 개인 식별 가능성, 목적 외 이용 등 프라이버시와 기업의 수익성 해결 중요

자료 활용을 위한 자발적인 민관 파트너십

- 기업 내부 자료를 이용한 통계를 작성 후 정부와 공유
- 기업이 자료를 제공하여 통계 당국이 작성
- 자료를 신뢰할 수 있는 제3자에 제공
- 통계기관의 자료 처리 기능을 민간 기업에 위탁

● 증거기반 정책수립(The Promise of Evidence-Based Policymaking), 2017

- 자료 보안과 프라이버시 보호를 강화하면서 행정 자료의 활용 가능 판단
- 자료 접근의 제한성, 부적절한 프라이버시 관행, 정책결정을 뒷받침하는 증거 생성 능력의 부족을 해결하기 위한 포괄적인 전략 제시
- **국립데이터서비스청(National Secure Data Service, NSDS) 신설**
 - * 통계 목적의 여러 기관 보유 자료의 획득, 결합 업무를 담당
 - * 자료보관소(data warehouse) 역할 지양 명시
- 일반 공개가 예정된 비식별화 기밀자료 대상 종합적 위험평가 실시
- 정책 결정에 활용될 기밀자료에 최신 프라이버시 강화 기술 적용
- 정부 부처 내 고위 공무원을 **자료관리 책임자로 지정**
- 정부 의사결정의 기준인 **연방통계의 무결성 및 객관성 유지 정책 성문화**
- 범정부적 증거구축 활동 조정 위한 대통령실 관리예산처의 부처간 조정 기능 활성화와 **조직 개편**
- **행정 절차를 증거구축 활동에 부합하도록 개선**

● 신산업구조비전(경제산업성, 2016)

대응 전략		구체적 대응 정책
1	데이터 활용촉진을 위한 환경정비	· 데이터 플랫폼 구축, 데이터 유통시장의 창설, 개인 데이터 이용·활용 촉진, 보안 기술 및 인력양성이 가능한 생태계 조성, 지적재산정책과 경쟁정책의 정비
2	전문인력 육성·획득, 고용시스템의 유연성 향상	· 새로운 수요에 대응한 교육시스템 구축, 글로벌 인재 획득, 다양한 노동참가 촉진, 노동시장, 고용시스템의 유연성 향상
3	이노베이션/기술개발의 가속화	· 오픈 이노베이션 시스템 구축, 핵심 이노베이션 거점 정비, 국가 프로젝트 구축, 지적재산관리와 국제표준화의 전략적 추진
4	재무기능 강화	· 리스크 머니(위험자금) 공급을 위한 equity finance 강화, 4차 산업혁명에 대응한 무형자산 투자 활성화, 핀테크 등 금융·결제 기능의 고도화
5	산업구조/취업구조 전환의 원활화	· 신속·과감한 의사결정이 가능한 거버넌스 체제 구축, 신속·유연한 사업재생/산업재편이 가능한 제도·환경 정비, 노동시장과 고용제도의 유연성 향상
6	중소기업/지방에 4차 산업혁명 파급	· 중소기업·지방에 IoT 등의 도입/이용 기반 구축
7	4차 산업혁명을 향한 경제사회시스템 구축	· 각종 규제 개혁, 행정서비스 향상, 전략적 제후를 통한 글로벌 전개 강화, 사회에 4차 산업혁명을 침투

자료: 사공목(2017), 일본의 4차 산업혁명 대응 실태와 시사점(KIET). 원 출처는 일본 경제산업성, 신산업구조비전(2016. 4)

● 관민데이터활용추진기본법, 2017

- 국가, 지방자치단체, 민간 기업이 보유한 데이터를 활용, 새로운 사업을 창출하고 이를 기반으로 국가 경쟁력 강화를 도모하기 위해 제정
 - * 온라인 행정절차의 원칙화
 - * 중앙정부와 지자체의 오픈데이터화
 - * 정부와 지자체의 데이터 활용 촉진 위한 시스템 규격 정비와 호환성 확보
 - * 디지털 격차 대책 수립

● 통계개혁추진회의최종종합, 2017

- 증거기반정책의 정착, 국민의 요구에 대응 등의 관점에서 과감한 통계개혁과 통계시스템의 정비를 추진
 - * 경제통계 개선, 이용자 관점의 통계시스템 재구축 및 활용 촉진
 - * 보고자 부담 경감 및 통계 업무·행정 체제의 재검토·업무의 효율화
- **통계 업무효율화를 위해 BPR(Business Process Reengineering) 실시**
- 통계 개혁 위한 **인재확보**와 **교육 지침 제정**, 추진
- 중앙 정부와 지방 통계기구의 업무 조정을 통한 체제 정비
- **통계위원회의 기능 강화**

● 미래투자전략2018(Society 5.0 데이터추동사회로의 혁신), 2018

- 4차 산업혁명으로 생활과 산업, 경제활동, 행정인프라, 지역·커뮤니티·중소 기업, 인재가 변화하는 새로운 사회(Society 5.0) 예견
- 행정서비스도 출발부터 완료까지 디지털화로 시간과 비용을 대폭 절감
- **정부 보유 데이터를 개방, 데이터를 활용한 혁신과 뉴비즈니스 창출 촉진**

Society 5.0 실현을 위한 행정 분야의 플래그십 프로젝트

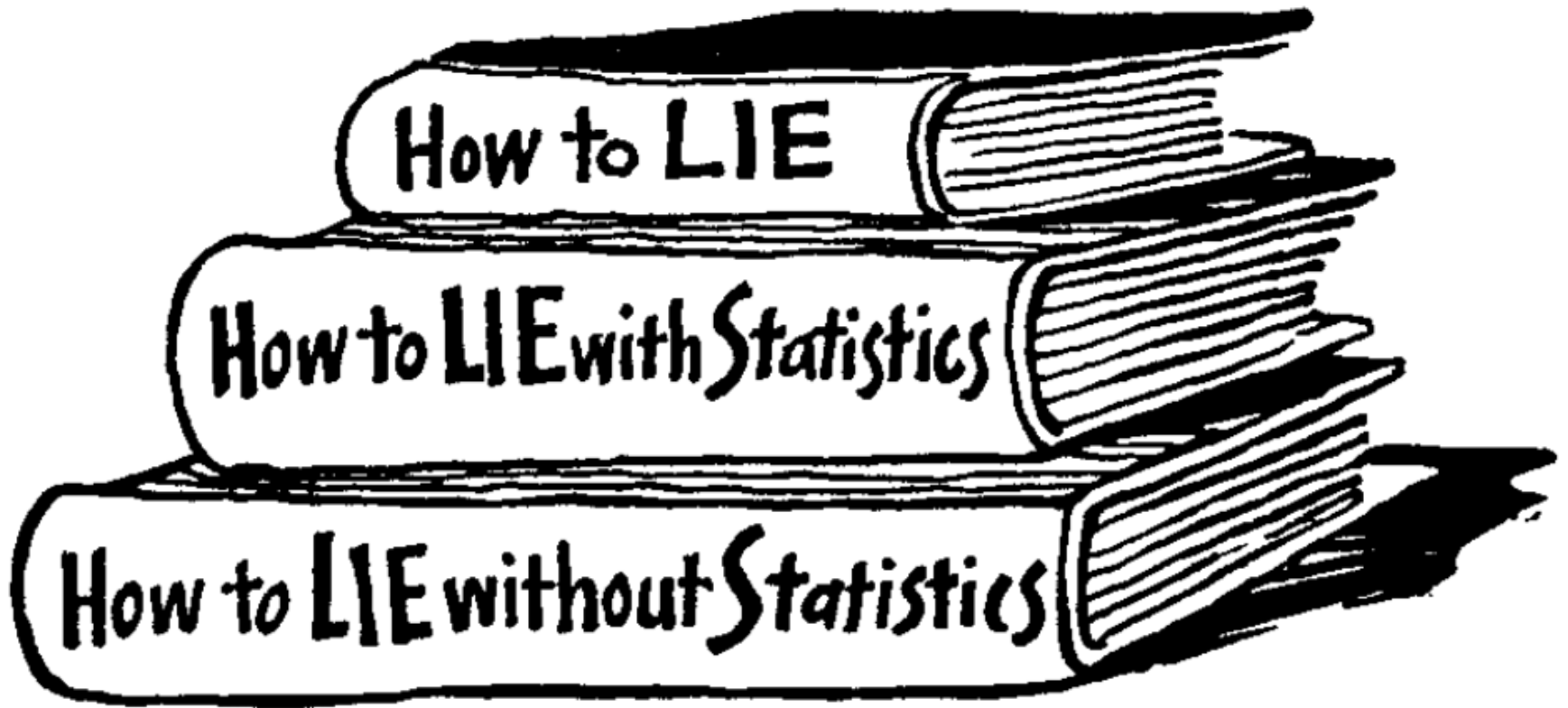
- 디지털정부의 추진: '디지털 first법' 국회 제출
행정서류 간소화, 날인/대면형 본인인증제도 개선, 지불수수료 온라인화, API 정비
민간이 필요로 하는 공공교통 데이터, 외국인의 내국 소비 데이터, 방재 데이터 조기 개방
- 차세대 인프라·메인テナンス·시스템 구축:
인프라와 연계된 모든 데이터를 집약, 공유하는 '인프라·데이터플랫폼' 구축
건설부터 유지관리까지 전체 프로세스 데이터를 클라우드화



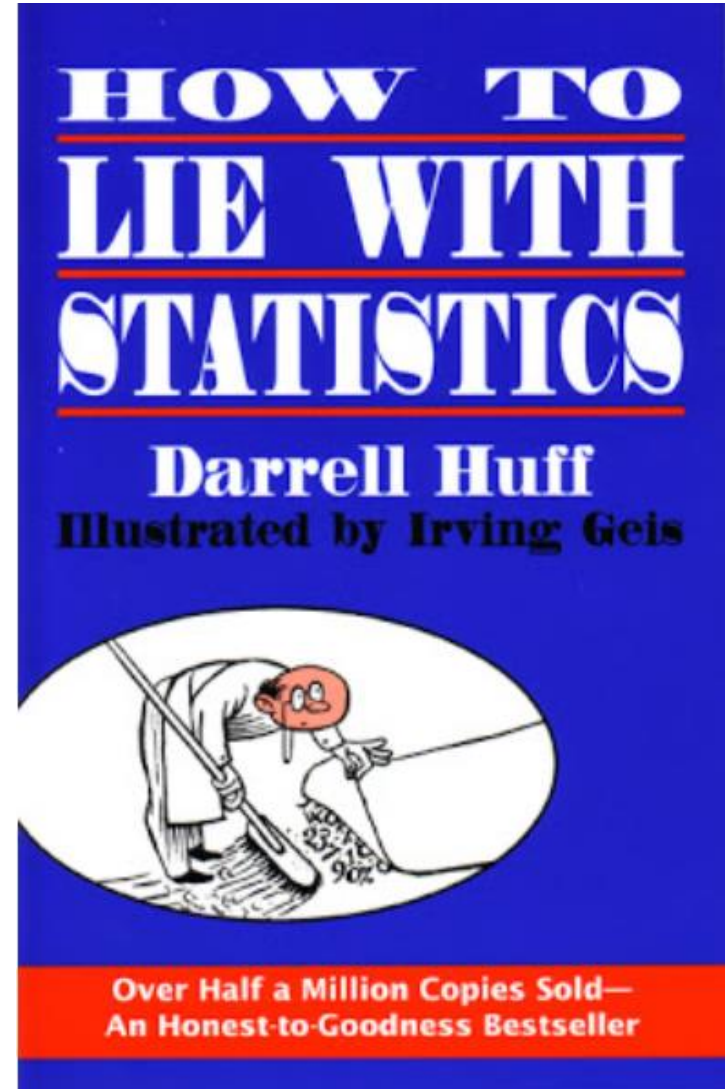
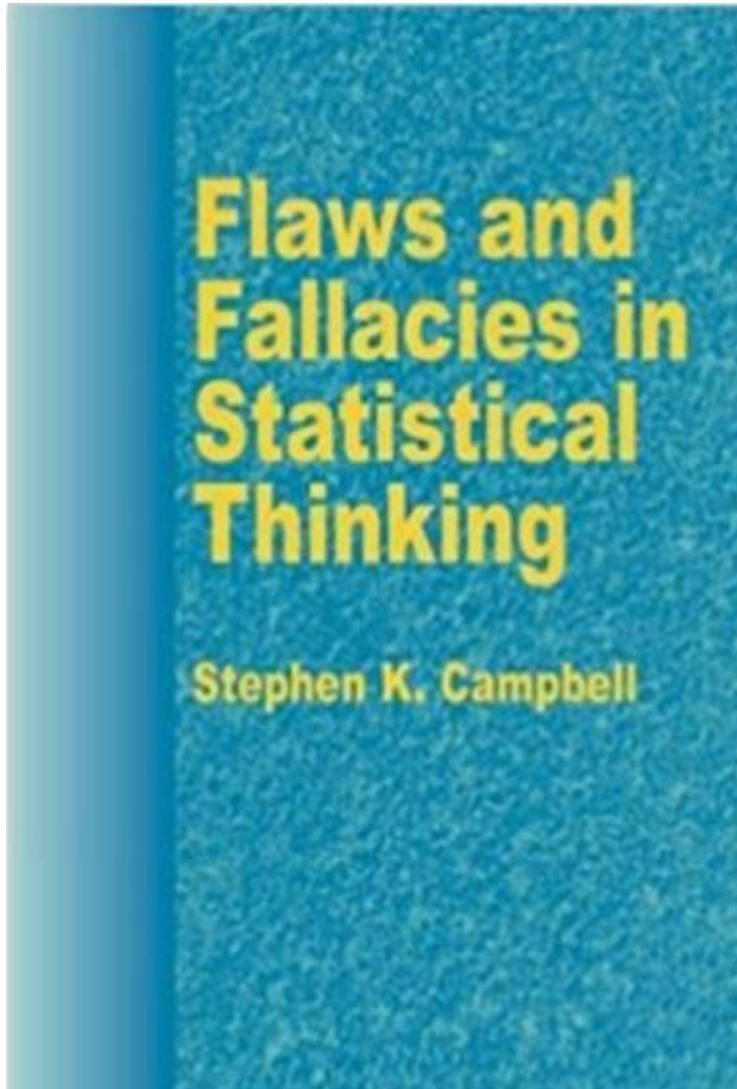
"That's what I want to say. See if you can find some statistics to prove it."



<http://matt dickenson.com/2012/08/06/statistics-ethics-and-open-data/>



<https://salyee.wordpress.com/2010/11/06/week-11-infographics/>

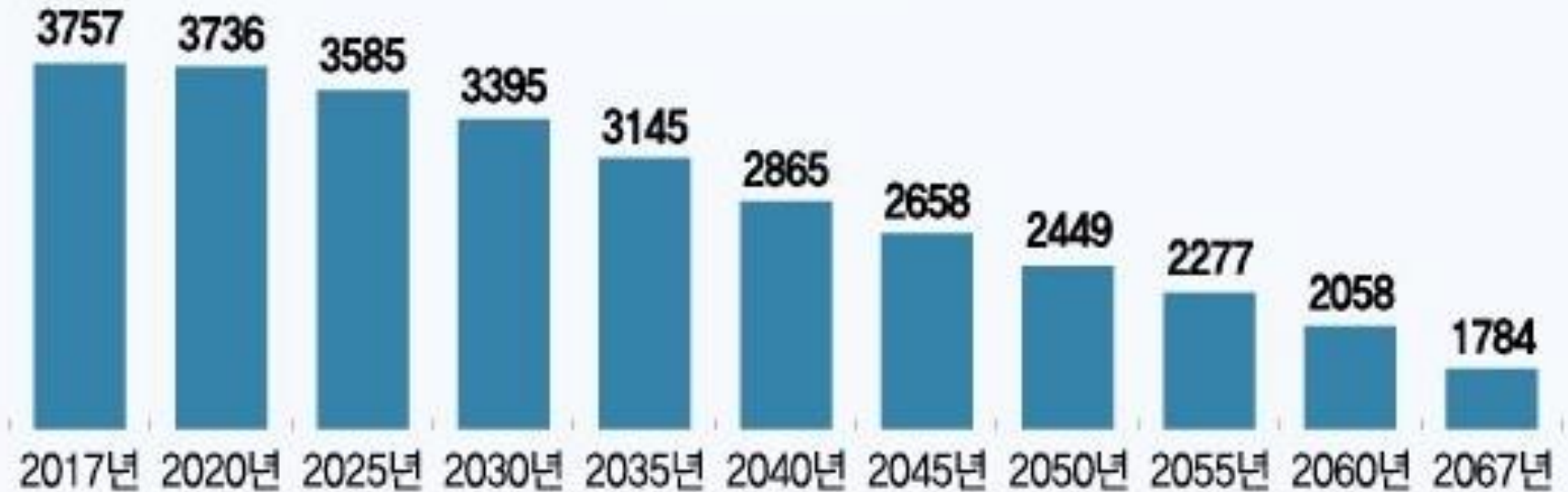


<https://www.amazon.com/>

출처: 정용찬(2018). '빅데이터의 통계적 활용과 데이터 윤리', 인터넷 윤리학회 춘계학술대회(2018. 5.25).

● [단독] 인구 절벽 ‘쓰나미’ ... 교육, 국방, 복지정책 확 바꾼다

생산가능인구 추계 *2017~2067년, 단위: 만명



*자료: 통계청 장래인구특별추계(중위추계)

그래픽: 유정수 디자인기자

<http://news.mt.co.kr/mtview.php?no=2019040216491257857&outlink=1&ref=http%3A%2F%2Fsearch.daum.net>

구분	데이터 유형	주요 내용
물가 통계	온라인 데이터 (U.S. Census Bureau)	표본조사 데이터와 온라인 거래 데이터의 연계 자동차부품, 보석/시계 POS데이터 분석 소지역 추정량 산출 모형 개발
인구 통계	모바일폰 데이터 (Eurostat, 벨기에통계청)	모바일폰 데이터와 상주인구통계 비교분석 인구, 관광, 교통, 이사 통계에 활용 가능
실업 통계	온라인 데이터 독일통계청	구인구직사이트((job portal)의 데이터를 활용, 일자리 추정
관광 통계	교통량/모바일폰/신용카드 데이터 스페인통계청	빅데이터와 조사 데이터간의 비교 분석 (국적, 관광 일수, 지출 규모 등)

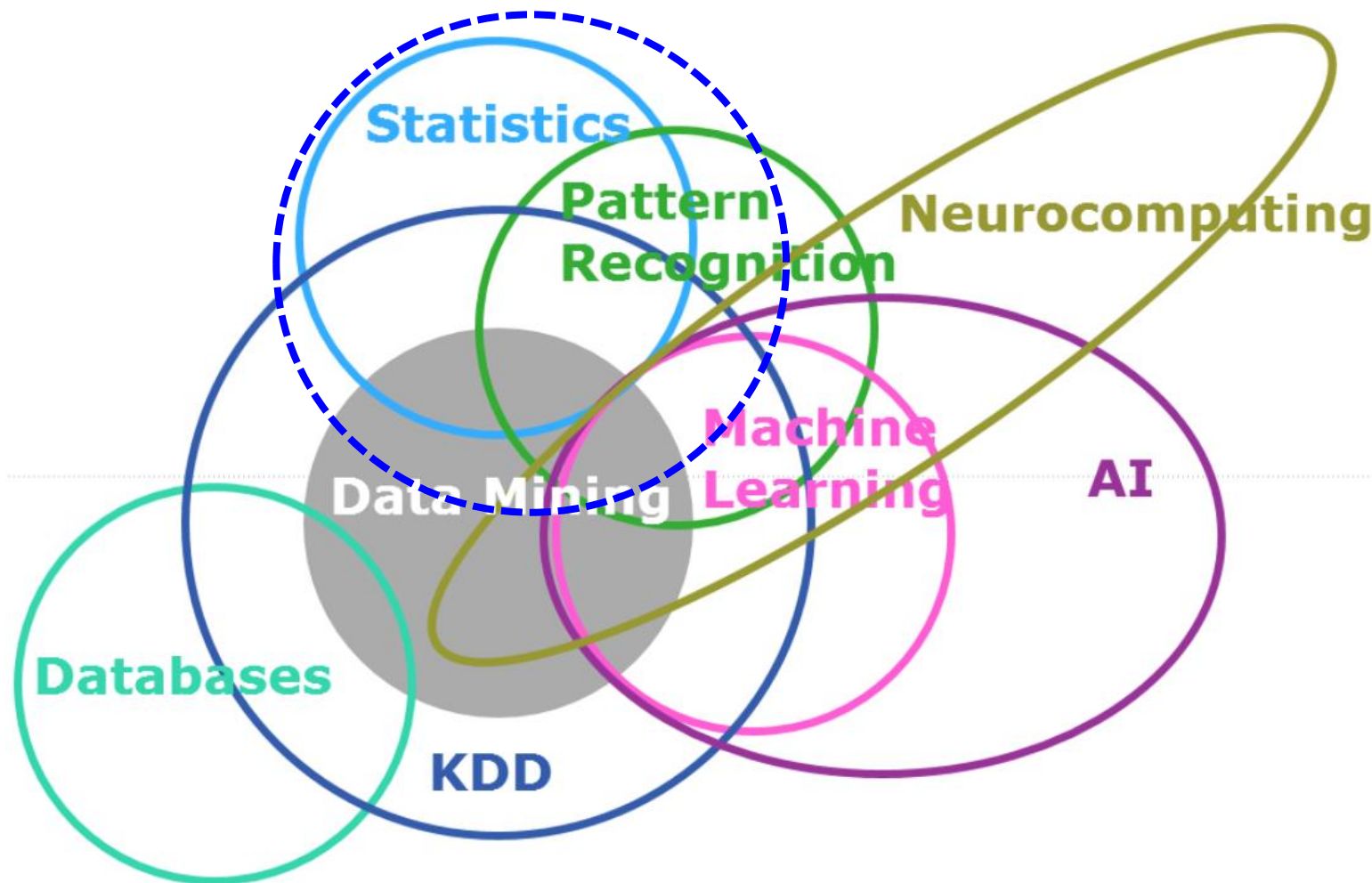
Source: Eurostat, Q2016 European Conference on Quality in Official Statistics, 31st May – 3rd June.
<http://www.q2016.es/scientific-information/conference-programme/>

구분	빅데이터	국가통계
관련성(relevance)	분야별 상이	높음
정확성(accuracy)	분야별 상이	높음
시의성(timeliness)	매우 높음	다소 낮음
생산 비용 (cost)	비교적 낮음	높음
응답부담(burden)	낮음	높음

출처: 정용찬(2016). '빅데이터의 국가통계 활용 전략', 통계청 국가통계발전포럼.

Statistical model vs Big Data model

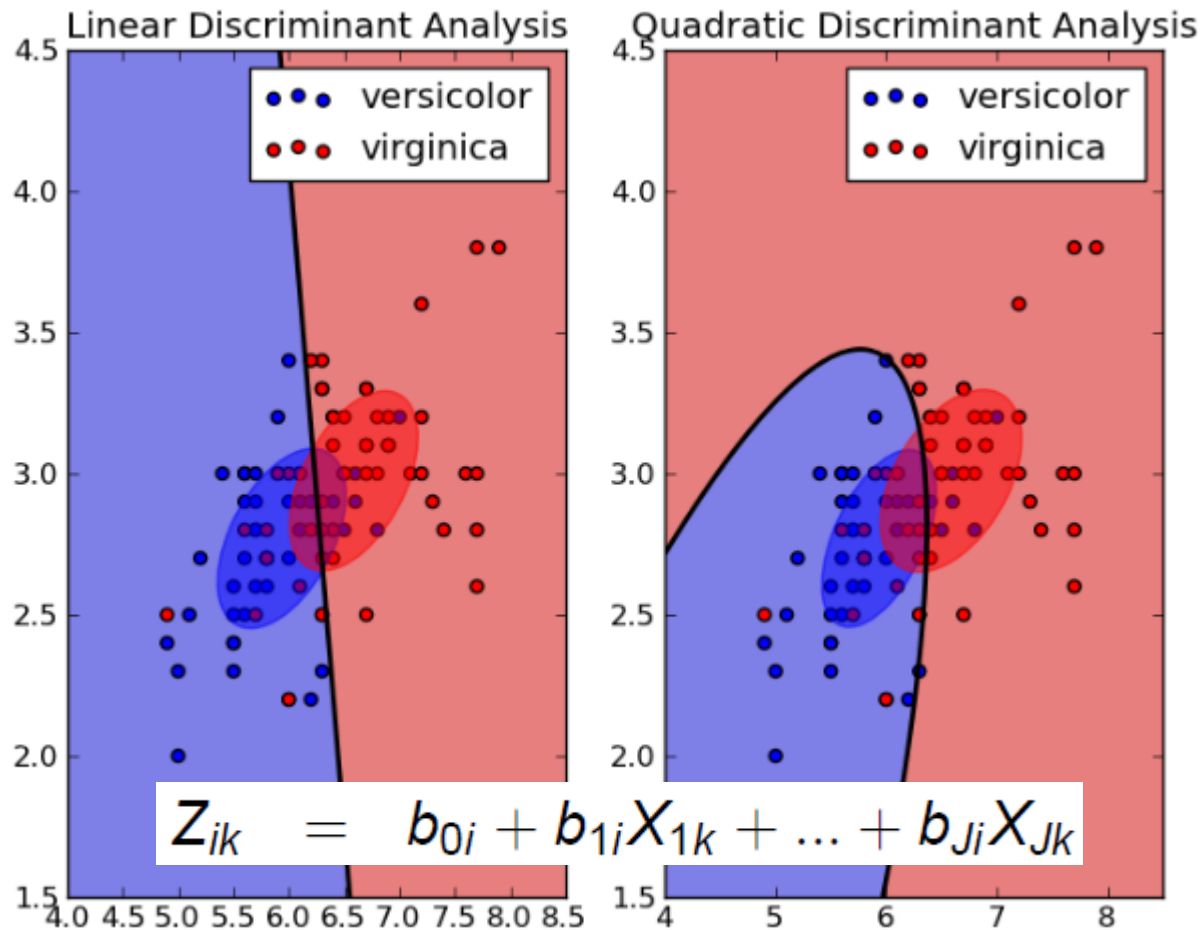
3. 데이터 시대의 통계



<https://www.analyticsvidhya.com/blog/2015/07/difference-machine-learning-statistical-modeling/>

Statistical model vs Big Data model

- discriminant analysis, factor analysis, decision tree...

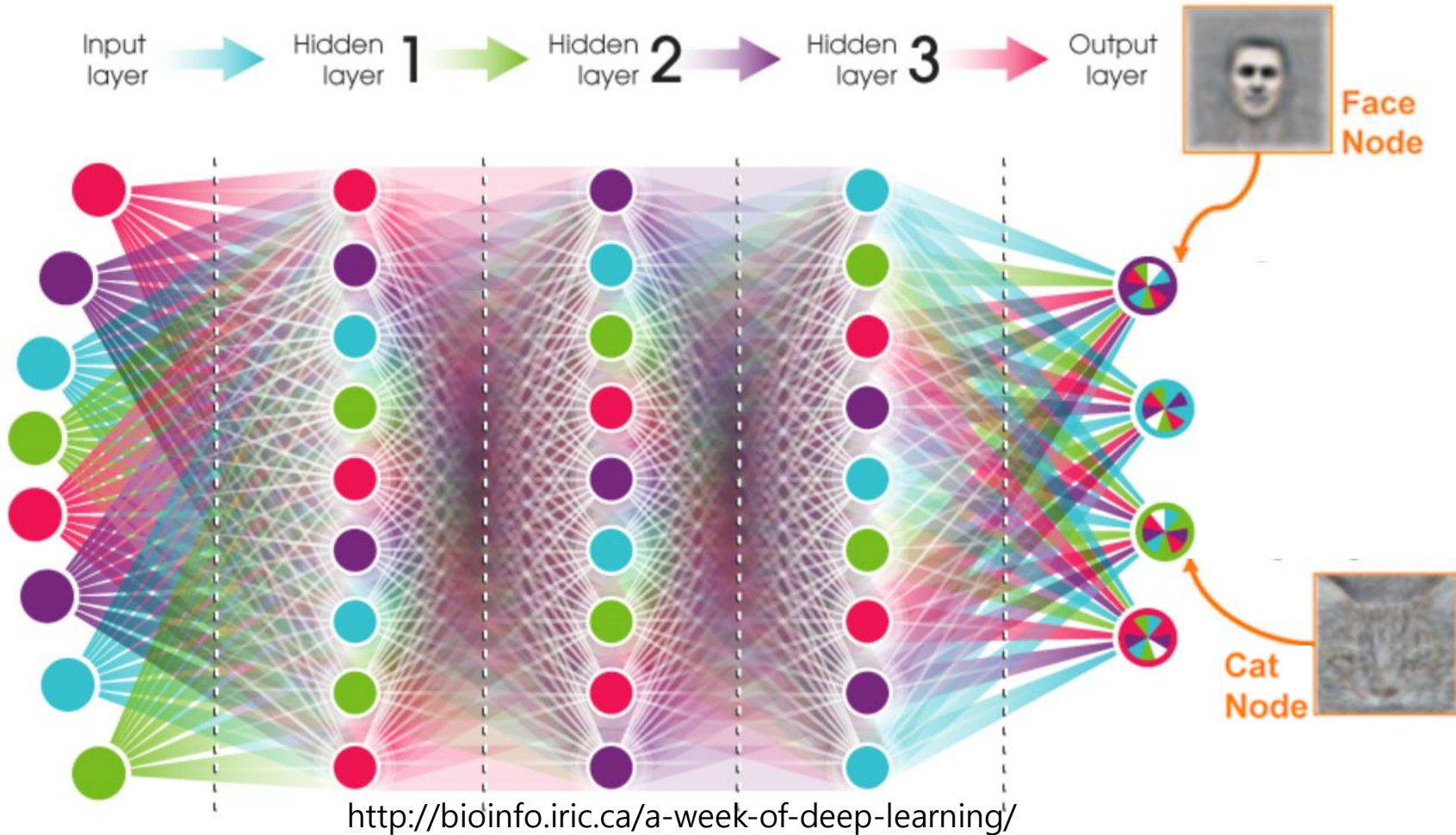


<https://ratsgo.github.io/machine%20learning/2017/03/21/LDA/>

Statistical model vs Big Data model

- machine learning, deep learning, artificial intelligence...

DEEP NEURAL NETWORK



국가 차원의 데이터전략 수립

통계청의 독립성과 조정 기능 강화

행정데이터와 통계생산의 유기적 연계

데이터 거버넌스와 통계 거버넌스 조화

소비자 행동 빅데이터 공유와 활용

지방분권 시대 준비 위한 데이터 거버넌스

개인정보 보호와 데이터 활용 위한 R&D

● 데이터 시대에 부합하는 정부 혁신 전략

- 데이터 생애주기(생산, 공유, 활용) 과정에서 **정부의 역할 재정립**
- 데이터 기반 정책결정 문화 정착을 위한 행정업무 재설계

● 데이터 거버넌스 강화를 위한 조직 전문화 전략

- 미국: 백악관의 통계수석 외에 수석 데이터과학자(Chief Data Scientist)
- 영국: 최고 데이터 책임자(Chief Data Officer) 신설

한국 통계법

제7조 ① 통계작성기관의 장은 통계의 작성 및 보급에 **필요한 인력과 예산을 확보하도록 노력**하여야 한다.

● 데이터 역량 강화 전략

- 초중등 교육 과정 개편(컴퓨팅과 데이터 분석 능력 강화)
- 공무원의 분석 능력 강화 정책 마련
- 기업의 데이터 분석 교육 지원 프로그램 개발
- * 영국의 Strategy for UK data capability, 2013

국가	통계 생산 체제	통계 조정 기구
영국	통계청과 주요 부처가 국가 통계를 생산하는 분산형 체제	의회를 지원하는 독립기구인 통계위원회가 통계청 감독, 통계 조정 기능 담당
미국	재무부(센서스국), 노동부(노동통계국) 등 주요 부처가 국가 통계를 생산하는 분산형	대통령실 관리예산처의 통계 수석이 통계조정 기능 담당
일본	총무성 통계국과 부처가 국가 통계를 생산하는 분산형	행정 관리를 총괄하는 총무성이 통계 조정 기능 담당
한국	기재부 통계청과 부처가 국가 통계를 생산하는 분산형 체제	국가통계위원회(위원장 기재부장관)가 통계 조정 기능 담당 통계청이 통계 조정 실무 담당

출처: 정용찬(2018). '4차 산업혁명 시대의 데이터 거버넌스 개선 방향'(KISDI Premium Report 18-05, 2018. 8. 7)

● 데이터 공유와 연계 위한 규제 개혁

- 행정 부처 간 자료 공유 활성화 위한 법과 행정 규제 개혁
- 중앙정부/지자체의 행정자료와 통계자료의 연계를 위한 제도 개선
- 행정자료 대체를 전제로 한 국가통계 생산체계의 개혁

* 영국 디지털경제법(2017): 부처간 통계 작성 목적의 정보공유 조항 신설

● 통계 생산 관점의 행정 업무 재설계

- 행정 업무 과정에서 필요 통계가 자연스럽게 생성, 축적될 수 있도록 BPR(Business Process Reengineering) 시행
- 응답자 부담을 고려

혼인신고서

※ 타인의 서명 또는 인장을 도용하여 허위의 신고서를 제출하거나, 허위신고를 하여 가족관계등록부에 실제와 다른 사실을 기록하게 하는 경우에는 **형법에 의하여 처벌**받을 수 있습니다. **눈표(*)로 표시한 자료**는 국가통계작성을 위해 통계청에서도 수집하고 있는 자료입니다.

※ 아래 사항은 「**통계법**」 제24조의2에 의하여 **통계청에서 실시하는 인구동향조사**입니다. 「통계법」 제32조 및 제33조에 의하여 성실응답의무가 있으며 개인의 비밀사항이 철저히 보호되므로 사실대로 기입하여 주시기 바랍니다.

※ 첨부서류 및 혼인당사자의 국적과 혼인종류는 국가통계작성을 위해 통계청에서도 수집하고 있는 자료입니다.

인구동향조사

		년		월		일부터 동거	
㉞ 실제 결혼 생활 시작일							
㉟ 혼인종류	남편	① 초혼 ② 사별 후 재혼 ③ 이혼 후 재혼		아내	① 초혼 ② 사별 후 재혼 ③ 이혼 후 재혼		
㊱ 최종 졸업학교	남편	① 학력 없음 ② 초등학교 ③ 중학교		아내	① 학력 없음 ② 초등학교 ③ 중학교		
	(부)	④ 고등학교 ⑤ 대학(교) ⑥ 대학원 이상		(처)	④ 고등학교 ⑤ 대학(교) ⑥ 대학원 이상		
㊲ 직업	남편 (부)	① 관리직 ② 전문직		아내 (처)	① 관리직 ② 전문직		
		③ 사무직 ④ 서비스직			③ 사무직 ④ 서비스직		
		⑤ 판매직 ⑥ 농림어업			⑤ 판매직 ⑥ 농림어업		
		⑦ 기능직 ⑧ 장치·기계 조작 및 조립			⑦ 기능직 ⑧ 장치·기계 조작 및 조립		
		⑨ 단순노무직 ⑩ 군인			⑨ 단순노무직 ⑩ 군인		
		㉑ 학생·가사·무직			㉑ 학생·가사·무직		

제24조의2(사법기관 등의 자료 제공) ① 통계청장은 통계의 작성을 위하여 필요한 경우에는 가족관계등록전산자료의 제공을 법원행정처장에게 요청할 수 있다.

② 통계청장은 통계의 작성을 위하여 필요한 경우에는 사망원인통계에 관련된 형사사법정보의 제공을 경찰청장 및 해양경찰청장 등에게 요청할 수 있다. <개정 2014. 11. 19., 2017. 7. 26.>

④ 통계청장은 다음 각 호의 자료를 활용하여 출생, 사망, 혼인, 이혼 등 인구 변화를 파악하기 위한 통계(이하 "인구 동태통계"라 한다)를 작성하여야 한다. <신설 2016. 12. 27.>

1. 제1항에 따른 가족관계등록전산자료
2. 다른 법률에 따라 출생, 사망, 혼인, 이혼 등의 신고를 하면서 신고인이 작성하여 제출한 자료 등 대통령령으로 정하는 자료

출생신고서

출생자에 관한 사항		
⑤ 임신주(週)수	임신 <input type="text"/> 주 <input type="text"/> 일	⑥ 신생아체중 <input type="text"/> . <input type="text"/> <input type="text"/> kg
⑦ 다태아 여부 및 출생순위	<input type="checkbox"/> 단태아 <input type="checkbox"/> 쌍태아(쌍둥이) <input type="checkbox"/> 삼태아(세쌍둥이) 이상	→ 쌍둥이 중 <input type="checkbox"/> 첫번째 <input type="checkbox"/> 두번째 → <input type="checkbox"/> 쌍둥이 중 <input type="checkbox"/> 번째
출생자의 부(父)에 관한 사항		출생자의 모(母)에 관한 사항
⑧ 실제생년월일	양력 / 음력 년 월 일	양력 / 음력 년 월 일
⑨ 최종졸업학교	<input type="checkbox"/> 무학 <input type="checkbox"/> 초등학교 <input type="checkbox"/> 중학교 <input type="checkbox"/> 고등학교 <input type="checkbox"/> 대학(교) <input type="checkbox"/> 대학원 이상	<input type="checkbox"/> 무학 <input type="checkbox"/> 초등학교 <input type="checkbox"/> 중학교 <input type="checkbox"/> 고등학교 <input type="checkbox"/> 대학(교) <input type="checkbox"/> 대학원 이상
⑩ 직 업	* 주된 일의 종류와 내용을 기입 합니다.	
⑪ 실제결혼생활시작일	년 월 일 부터	
⑫ 모의 총출산아 수	이 아이까지 총 <input type="text"/> 명 출산 (<input type="text"/> 명 생존, <input type="text"/> 명 사망)	

②란: 부(父)에 관한 사항: 혼인외 출생자를 모(母)가 신고하는 경우에는 기재하지 않으며, 전혼 해소 후 100일 이내에 재혼한 여자가 재혼성립 후 200일 이후, 직전 혼인의 종료 후 300일 이내에 출산하여 모가 출생신고를 하는 경우에는 부의 성명란에 "부미정"으로 기재합니다.

등록 기준지: 부모가 모두 외국인이면 그 부모의 국적(출생신고 당시)을 각각 기재하며, 부모 중 한쪽이 외국인이면 외국인인 부 또는 모의 국적을 기재합니다.

⑩란: 직업: 아이가 출생할 당시의 부모의 주된 직업을 말하며 주된 일의 종류와 내용을 사업장명과 함께 기재합니다.

<예시> ○○회사 영업부 판촉사원, ○○상가에서 의류판매, 우리 논에서 논농사

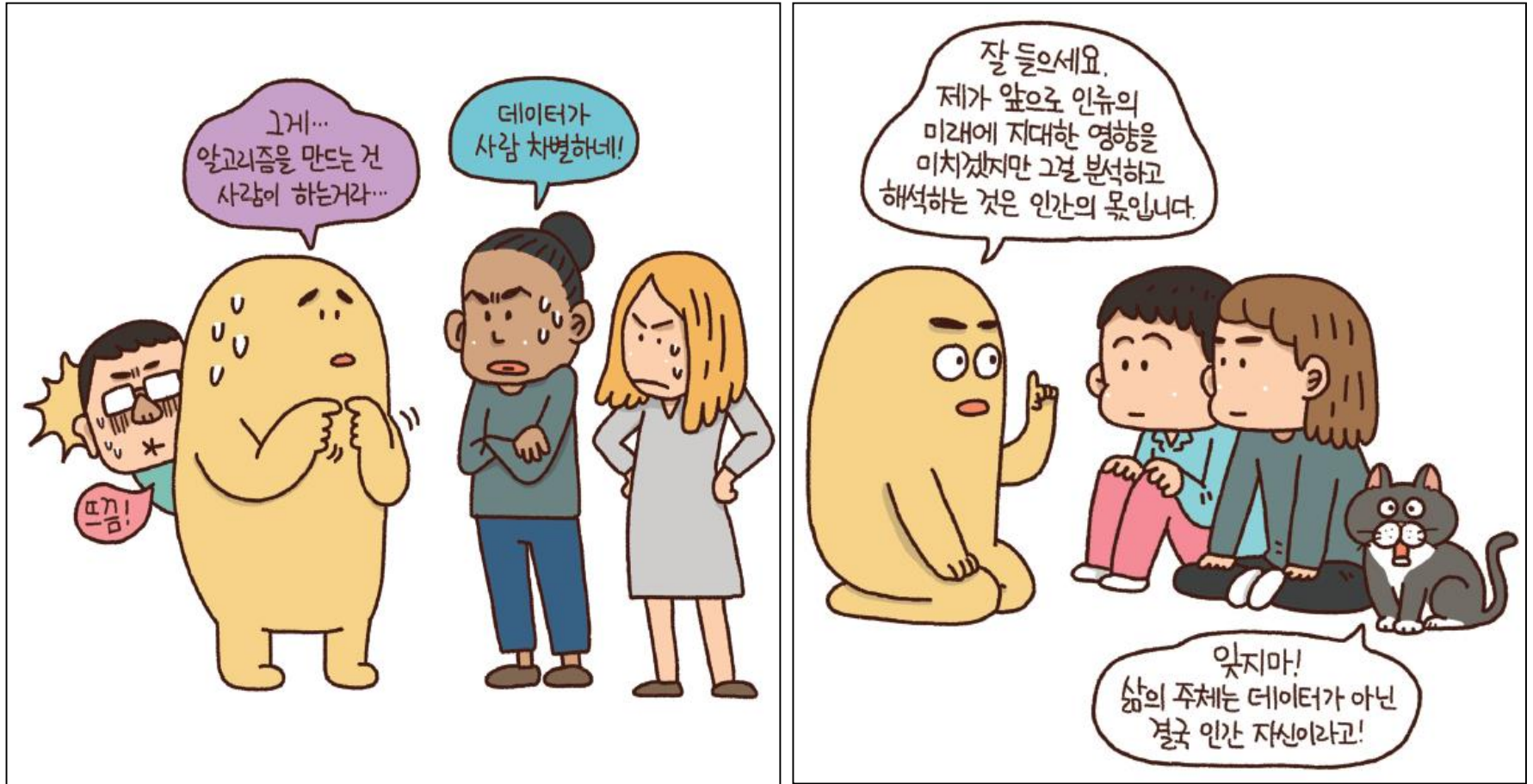
⑫란: 모의 총 출산아수: 신고서상의 본 아이까지 모두 몇 명의 아이를 출산했고 그 중 생존아와 사망아 수를 기재하며, 모가 재혼인 경우에는 이전의 혼인에서 낳은 자녀까지 포함합니다.

구분	행안부	통계청
위원회	공공데이터전략위원회(총리) 데이터기반행정 활성화 위원회(위원장 행안부장관 지명)	국가통계위원회(기재부장관)
법제	공공데이터의 제공 및 이용 활성화에 관한 법률 데이터기반행정 활성화에 관한 법률안 (‘17. 12. 제출) - 데이터제공 요청(제11조) - 민간데이터 제공 요청(제15조) - 데이터기반행정책임관(제20조)	통계법 - 행정자료 활용 통계작성(제18조) - 행정자료의 제공(제24조) - 통계기반정책평가 (제12조2) - 통계책임관 지정 및 운영(제6조)
인프라	공공데이터포털(data.go.kr) 공공빅데이터센터 구축계획 발표(‘18. 1)	국가통계포털(kosis.kr) 운영 통계빅데이터부산센터 개소(‘18. 6), 대전, 서울 개소 예정(‘19)

● 데이터 슈퍼파워의 과거와 미래



**“인터넷을 가장 잘 다루는 나라에서
데이터를 가장 잘 다루는 나라로”**



출처: 정용찬, 생각이 크는 인문학 15. 빅데이터, 2018, 을파소.

“경청해 주셔서 감사합니다”



- 정용찬, KISDI ICT통계정보연구실 데이터사이언스 그룹장. ycjung@kisdi.re.kr
- 저서 : 생각이 크는 인문학 15. 빅데이터(2018), 빅데이터(2013), 시청률조사(2013), 미디어산업통계(2013), 미디어소비통계(2013), 융합과 통섭(2012, 공저), 미디어통계의 이해(2008), 통계로 본 대한민국(2007) 등
- 연구보고서: 정책지원 강화를 위한 국가통계 관리체계 개선 심층연구(2018), 4차 산업혁명 시대의 데이터 거버넌스 개선 방향(KISDI Premium Report 18-05, 2018. 8. 7), 4차 산업혁명 시대의 데이터 경제 활성화 전략(KISDI Premium Report 17-04, 2017. 6. 14), 조사환경변화에 대응한 ICT총계 생산체계 혁신방안연구(KISDI, 2017), 빅데이터 산업과 데이터브로커(2015, KISDI Premium Report), 통신산업의 빅데이터생태계 촉진 전략(2015), ICT통계관리체계 개선방안 연구(2015), 빅데이터산업 촉진전략연구(2014) 등