

디지털융합훈련과정

빅데이터는 나의 친구

나도 빅데이터 전문가

학습자용 학습자료



과정핵심학습자료



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
2 차시	빅데이터로 변화되는 우리의 일상
학습자료	<p>빅데이터의 역기능</p> <ol style="list-style-type: none">1. 개인이 인터넷 또는 모바일에 올린 모든 대화들을 조합하면 새로운 정보가 재생성될 수 있으며,2. 이렇게 수집된 비식별화된 데이터들이 몇 번의 과정을 거쳐 분석 · 처리되면 실제 개인을 식별할 수 있는 정보로 재탄생할 수 있고,3. 일례로 미국의 한 대형 소매 유통업체가 고객이 임신부용 속옷과 튜브 크림 등을 구매했던 내역을 바탕으로 고객의 임신 사실을 알아내어 고객에게 임신부용 쿠폰을 보낸 사례가 있으며,4. 2013년 미국 국가안보국이 NSA가 미국인들의 수백 만 건의 통화 기록과 인터넷 데이터 등의 개인정보를 무차별적으로 수집하고 감시하고 있다는 내용이 폭로되었으며,5. 카카오톡이 국정원의 사이버 검열에 협조하여, 국민들이 러시아의 메신저인 텔레그램으로 사이버 망명을 한 사례가 있다.



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
3차시	빅데이터가 뭐길래 난리야
학습자료	<p>빅데이터의 특징</p> <ol style="list-style-type: none">1. 데이터 크기 단순 저장되는 물리적 데이터양을 나타내며, 빅 데이터를 사용하면 저밀도 비정형 데이터를 대량으로 처리해야 합니다.2. 데이터 속도 데이터가 얼마나 빨리 수신 및 처리되는가를 말합니다.3. 데이터 종류(다양성) 사용 가능한 데이터의 유형 수를 나타내며 비정형 및 반정형 데이터 유형은 의미를 도출하고 메타 데이터를 지원하기 위해 추가로 전처리가 필요합니다.4. 정확성 빅데이터를 분석하는 데 있어 기업이나 기관에 수집한 데이터가 정확한 것인지 분석할 만한 가치가 있는지 등을 살펴보는 것을 말합니다.5. 가변성 실제로 자신의 의도와는 다르게 의미가 전달될 수 있으므로 빅데이터의 새로운 속성으로 가변성(Variability)이 제시되고 있습니다.6. 시각화 - 정보의 사용 대상자가 쉽게 이해할 수 있어야 하므로 데이터를 시각화할 수 있어야 합니다.



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
3 차시	빅데이터가 뭐길래 난리야?
학습자료	<p>건강보험심사평가원 건강보험 빅데이터 개방 사례</p> <ol style="list-style-type: none">1. 2020년 초, 세계 곳곳의 연구자들이 백신을 개발하기 위해 모든 노력을 하였지만 '임상 데이터' 부족으로 결론에는 도달하지 못하고 있었다.2. 건강보험심사평가원이 전 세계 연구자들에게 의료 빅데이터를 개방하여 '코로나19 국제협력 연구'가 시작되면서 코로나19 연구에 속도가 나기 시작하였다고 한다.3. 건강보험심사평가원이 전 국민 의료데이터를 개방하고 국제협력 연구를 시작할 수 있었던 배경에는, 단일보험체제인 우리나라 건강보험 제도적 특성이 작용했다.4. 건강보험심사평가원은 '코로나19 국제협력 연구' 성과를 공유하고 향후 발전방안을 논의하기 위해 국내외 전문가들이 참여한 'HIRA 빅데이터 미래 포럼'을 개최하여 유튜브를 통해 생중계하였다.5. 코로나19 환자의 임상 데이터 부족으로 연구에 어려움을 호소하던 전 세계 연구자들의 관심이 높아졌고 이는 연구성과로 이어졌다.



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
3 차시	빅데이터가 뭐길래 난리야?
학습자료	<p>전통적인 데이터와 빅데이터 비교 요소</p> <ol style="list-style-type: none">1. 데이터의 크기2. 데이터 구성 방법3. 데이터 관리에 필요한 아키텍처4. 데이터가 파생되는 소스5. 데이터 분석에 사용되는 방법



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
3 차시	빅데이터가 뭐길래 난리야?
학습자료	<p>기존 데이터와 빅데이터 비교</p> <ol style="list-style-type: none">1. 크기 _ 기존 데이터는 기가바이트와 테라바이트 단위 측정, 빅데이터는 일반적으로 페타바이트, 제타바이트 또는 엑사바이트 단위로 측정한다.2. 구성 _ 기존 데이터는 정형 데이터로 SQL, Oracle DB 및 MySQL과 같은 사전 구성된 데이터베이스를 사용하며, 빅데이터는 다이내믹 스키마를 사용하기 때문에 정형화되어 있지 않은 원시 데이터 처리가 가능하다.3. 아키텍처 _ 기존 데이터는 중앙 집중식 아키텍처를 통해 관리되며, 빅데이터는 규모와 복잡성 때문에 분산 아키텍처를 필요로 한다.4. 데이터 출처 _ 기존 데이터는 ERP, CRM, 온라인 트랜잭션 및 기타 엔터프라이즈 레벨 데이터에서 파생되며, 빅데이터는 소셜 미디어, 디바이스 및 센서 데이터, 시청각 데이터 등의 텍스트, 동영상, 이미지 및 오디오 파일도 포함한다.5. 분석 _ 기존 데이터 분석은 기업들이 정해진 기간 동안 특정 전략이나 변경 사항이 제한된 범위의 영향을 이해하는 데 도움이 되며, 빅데이터는 초 단위로 생성되므로 데이터가 수집되는 동안 분석할 수 있습니다.



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
3 차시	빅데이터가 뭐길래 난리야?
학습자료	<p>기존 데이터와 빅데이터 비교</p> <ol style="list-style-type: none">1. 크기 _ 기존 데이터는 기가바이트와 테라바이트 단위 측정, 빅데이터는 일반적으로 페타바이트, 제타바이트 또는 엑사바이트 단위로 측정한다.2. 구성 _ 기존 데이터는 정형 데이터로 SQL, Oracle DB 및 MySQL과 같은 사전 구성된 데이터베이스를 사용하며, 빅데이터는 다이내믹 스키마를 사용하기 때문에 정형화되어 있지 않은 원시 데이터 처리가 가능하다.3. 아키텍처 _ 기존 데이터는 중앙 집중식 아키텍처를 통해 관리되며, 빅데이터는 규모와 복잡성 때문에 분산 아키텍처를 필요로 한다.4. 데이터 출처 _ 기존 데이터는 ERP, CRM, 온라인 트랜잭션 및 기타 엔터프라이즈 레벨 데이터에서 파생되며, 빅데이터는 소셜 미디어, 디바이스 및 센서 데이터, 시청각 데이터 등의 텍스트, 동영상, 이미지 및 오디오 파일도 포함한다.5. 분석 _ 기존 데이터 분석은 기업들이 정해진 기간 동안 특정 전략이나 변경 사항이 제한된 범위의 영향을 이해하는 데 도움이 되며, <u>빅데이터는 초 단위로 생성되므로 데이터가 수집되는 동안 분석할 수 있습니다.</u>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
3 차시	빅데이터가 뭐길래 난리야?
학습자료	<p>정형데이터와 비정형데이터 비교</p> <p>1. 정형 데이터</p> <ul style="list-style-type: none">- 보유하고 있는 정보에 대한 적절한 정의가 있는 경우와,- 구글 스프레드시트 또는 마이크로소프트 엑셀과 같은 스프레드시트 프로그램에 표시할 수 있는 모든 데이터를 말하며,- <u>일반적으로 관계형 데이터베이스라고 하며, SQL을 사용하여 작성, 읽기 및 조작 가능하다.</u>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
3 차시	빅데이터가 뭐길래 난리야?
학습자료	<p>정형데이터와 비정형데이터 비교</p> <p>1. 비정형 데이터</p> <ul style="list-style-type: none">- 정형 데이터가 아닌 모든 데이터는 비정형 데이터로 분류하고,- 2025년까지 우리가 접하는 데이터의 80%가 텍스트, 오디오, 이미지 또는 동영상 형태의 비정형 데이터가 될 것으로 예상되며,- 비정형 데이터에 가장 일반적으로 사용되는 데이터베이스는 SQL 데이터베이스의 기능을 넘어 더 넓은 범위의 데이터를 처리할 수 있음을 나타내는 NoSQL을 사용한다.



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
3 차시	빅데이터가 뭐길래 난리야?
학습자료	<p>가트너_ 빅데이터 특징 3V</p> <ol style="list-style-type: none">1. 빅데이터 특징 Gartner 3V _ Volume → 단순 저장 물리적 데이터양, 저밀도 비정형데이터 대량 처리2. 빅데이터 특징 Gartner 3V _ Velocity → 데이터 수신 및 처리 속도, 메모리로 직접 스트리밍 시 가장 빠른 속도3. 빅데이터 특징 Gartner 3V _ Variety → 사용 가능한데이터의 유형 수, 비정형 및 반정형 데이터 유형



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
4 차시	빅데이터는 어떻게 등장했지?
학습자료	<p>인류의 데이터 기술 발전</p> <ol style="list-style-type: none">1. 데이터 기술 발전 _ '수'의 발견 → 인류는 4만 년 전 '수' 발견 _ 일정한 계산 형식 필요 → '수'의 발견 _ 인류를 지구에서 가장 우수한 생물 종 변화 사건2. 데이터 기술 발전 _ '수'의 발견 → 원시 인류 _ 손가락 표시 체계 발전, 수를 세기 시작 추측3. 데이터 기술 발전 _ 엄대 / tally stick → 나무 막대기나 뿔, 뼈 등에 빗금으로 새겨서 표시 → 일종의 출납 장부, 원시적 계산을 통해서 식량 재고 파악4. 데이터 기술 발전 _ 문자 형태의 숫자 체계 → <u>기원전 4000년경 이란에서 처음 사용되었을 것으로 추정</u> → 점토판 사용 각 상품 마다 다른 숫자 표기



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
4 차시	빅데이터는 어떻게 등장했지?
학습자료	<p>1926년 / 니콜라 테슬라</p> <p>"무선통신이 완벽하게 적용되면 지구 전체가 거대한 두뇌로 전환될 것이다. 우리는 거리에 관계 없이 즉시 서로 통신할 수 있을 것이다. 그뿐 아니라, 텔레비전과 전화를 통해 우리는 수천 마일의 중간 거리에도 불구하고 서로 얼굴을 마주 보는 것처럼 완벽하게 보고 들을 수 있을 것이다."</p> <p>"이걸 할 수 있는 기계는 현재의 전화기에 비해서 놀라울 정도로 단순할 것이다. 조끼 주머니에도 넣고 다닐 수 있을 것이다. 우리는 대통령의 취임, 월드 시리즈 경기, 지진의 대혼란, 전쟁의 공포와 같은 사건들을 현장에 있는 것처럼 보고 들을 수 있게 될 것이다."</p>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
4 차시	빅데이터는 어떻게 등장했지?
학습자료	<p>빅데이터 등장 배경</p> <p><u>빅데이터 시대 개막 _ 소셜 데이터 혁명</u></p> <ul style="list-style-type: none">→ 사용자 자발적 정보 생성, 1990년대 정보 혁명과는 상이→ 소셜 네트워크 서비스 정보 _ 감성적 정보 큰 비중 차지→ 진실성, 진정성, 관련성 증가 _ 데이터로서 매우 높은 가치



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
4 차시	빅데이터는 어떻게 등장했지?
학습자료	<p>기업과 빅데이터</p> <ol style="list-style-type: none">1. 가치 있는 정보 취득 기대 증가 → 비정형데이터 _ 모바일과 SNS 등 바탕으로 폭발적 증가 _ 고객 욕구와 기호 파악, 새로운 수익창출 기대2. 경제 및 경영 환경의 불확실성 심화 → 데이터 신속 분석 _ 최적의 전략 선택 → 직관적 주먹구구식 의사결정 _ 발생 가능 막대한 피해 예방3. 차별화 전략 창출 → 빅데이터 기반 경영전략, 기업성과 및 성장지속 원동력



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
5 차시	국가 데이터 정책_대한민국 데이터 119 프로젝트
학습자료	<p>1. 미국 _ 선제적 전략투자, 세계 최고 경쟁력 보유</p> <ul style="list-style-type: none">→ 2012년 '빅데이터 R&D 이니셔티브' 발족→ 2016년 정부 빅데이터 연구개발 전략계획 발표→ '빅데이터 R&D 이니셔티브' _ 차세대 데이터과학자 양성 목적→ 정부 빅데이터 연구개발 전략계획 _ 기능 강조 및 정책 확장 목적 <p>2. 중국 _ 2017년 '디지털 중국' 건설 가속화 전략 발표</p> <ul style="list-style-type: none">→ 디지털 경제 _ 현대화 경제체계 내 빅데이터 발전 및 응용→ 사회 안정 _ 빅데이터 활용 사회보장 및 민생 개선 도모→ 데이터 거버넌스 _ 빅데이터 확보 및 안전중심 거버넌스 확충 <p>3. EU 2014 EC _ 데이터 주도 경제 결의안 채택</p> <ul style="list-style-type: none">→ 2014. EC-Big Data Value Association MOU 체결→ 2015. Digital Single Market Strategy for Europe→ 2017. Building an European Data Economy→ 2018. Toward a Common European Data Space→ 2018. 민간부문의 데이터 공유에 관한 가이드→ 2018. GDPR: General Data Protection Regulation→ 2018.Regulation on a framework for the free flow of non-personal data in the EU→ 2019 Open Data Directive: Directive on open data and the re-use of public sector information→ 2020. A European strategy for data



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
5 차시	국가 데이터 정책_대한민국 데이터 119 프로젝트
학습자료	<p>국가데이터 정책_기본 방향</p> <ol style="list-style-type: none">1. 공급자인 정부 중심이 아닌 기업과 시장 중심2. 국민을 배려하는 사람과 서비스 중심,3. 개별 정책보다는 체계적인 거버넌스 중심으로 설정하고,4. 민간의 오랜 요구에도 불구하고 해결이 어려웠던 11대 실천과제와 국민이 데이터 활용성과를 직접 체감할 수 있는 9대 서비스를 도출



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
5 차시	국가 데이터 정책_대한민국 데이터 119 프로젝트
학습자료	<p>국가데이터 정책_ 주요 내용</p> <ol style="list-style-type: none">1. 정부에서 그동안 개방하지 않았던 핵심데이터 즉시 제공2. 국민의 데이터 주권 강화를 위해 마이데이터 이용 활성화3. 기업이 데이터를 활용하여 산업경쟁력을 강화하도록 지원4. 특별 현안 과제로 코로나19 타임캡슐 프로젝트 시행5. 장기적으로 범 국가차원의 데이터 관리전략 수립



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
6 차시	빅데이터가 꿈꾸는 미래_스마트시티를 중심으로_1부
학습자료	<p>'스마트시티' 정의</p> <ol style="list-style-type: none">1. "첨단 정보통신 기술(CT)과 빅데이터 등 신기술을 접목해 각종 도시 문제를 해결하고 삶의 질을 개선할 수 있는 도시 모델"2. "경제, 이동성, 환경, 인간, 생활, 행정 등 다양한 주요 분야에서 우수하고 지속 가능한 경제 발전과 높은 삶의 질을 창출하는 발전된 도시"3. "첨단 정보통신 기술을 이용해 도시의 모든 인프라를 네트워크화한 미래형 첨단 도시"



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
6 차시	빅데이터가 꿈꾸는 미래_스마트시티를 중심으로_1부
학습자료	<p>스마트시티</p> <ol style="list-style-type: none">1. 빅데이터와 인공지능, 사물인터넷IoT, 5G 등 모든 기술이 융·복합 되어야 실현 가능하다.2. 디지털 트랜스포메이션 시대에 가장 기대되는 분야이다.3. 다양한 규모의 도시 교통, 에너지 효율성, 정부서비스 및 의료 서비스와 같은 핵심 분야에 큰 변화를 불러올 것이다. 다.



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
6 차시	빅데이터가 꿈꾸는 미래_스마트시티를 중심으로_1부
학습자료	<p>스마트시티 _ 에너지분야</p> <p>1. 건물에너지관리시스템 : BEMS → <u>빌딩 내 에너지 관리설비 정보 수집 분석 _ 효율적 에너지 사용</u> → <u>센서 및 계측장비 설치 _ 실시간 모니터링으로 제어</u></p> <p>2. 스마트그리드 : Smart Grid → 원격검침, 풍력발전 감시제어, 전기차 충전, 가정용 충전기 등 → 전기차 충전서비스 _ 통합관제서비스 활용, 탄소절감, 에너지 고효율화 추구 → 전기충전 표준화시스템</p> <p>3. 지능형 가로등 → 가로등 점멸기 _ 태양열 LED 조명등, 무선 인터넷, CCTV, 환경 센서 → 태양광 발전 가로등 _ 에너지 절약, 조명 밝기 연출 가능</p>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
6 차시	빅데이터가 꿈꾸는 미래_스마트시티를 중심으로_1부
학습자료	<p>스마트시티 _ 환경분야</p> <p>1. 스마트 워터그리드 → 침수, 홍수, 수질오염원 관리, 빗물재 활용, 물 수급 통합정보관리 등 → 도시침수와 홍수관리 _ 센서수집, 빅데이터 분석, 수위 및 침수 예상 → 하천 물 흐름 고려 _ 정보 제공으로 자연재해로 인한 피해 최소화</p> <p>2. 스마트 수목관리 → RFID, NFC 등 적용 _ 수목 생태정보 전송 및 관리시스템 → 수목의 이력관리, 병충해관리, 산불감시 등 활용</p> <p>3. 스마트 대기오염관리 → 탄소배출량 관리시스템 _ 대기오염물질 배출 최소화 시스템 → 오염원 분석, 예측결과 제공 _ 도시 발생 대기오염 배출정보 제공</p> <p>4. <u>스마트 쓰레기 관리</u> → <u>쓰레기통 센서 측정 쓰레기 처리 _ 청소업무 효율성 증가</u></p>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
6 차시	빅데이터가 꿈꾸는 미래_스마트시티를 중심으로_1부
학습자료	<p>스마트시티 _ 방법 · 안전 분야</p> <p>1. 지능형 방범서비스 → 영상분석기술 활용 _ 차량번호판, 안면, 홍채 등 인식 CCTV 설치 → 시민 안전과 효율적 범죄예방, 긴급상황에 대한 대처 등 가능 → <u>어린이 보호구역, 우범 지역, 골목길, 공원 등 설치 _ 효과 극대화</u></p> <p>2. 스마트 재난안전 시스템 → 재난문자방송 서비스 / CBS _ 재난사항 신속 전달 서비스 → 메시지 _ 국민안전처 수립 송출기준에 따라 송출</p>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
6 차시	빅데이터가 꿈꾸는 미래_스마트시티를 중심으로_1부
학습자료	<p>글로벌 스마트시티 구축 동향 _ Smart City Index</p> <p>2020년 기준 전세계에서 가장 스마트한 도시로, 싱가포르(1위), 헬싱키(2위), 취리히(3위), 오클랜드(4위), 오슬로(5위) 순이며, 부산과 서울은 각각 46위와 47위를 기록했다.</p>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
6 차시	빅데이터가 꿈꾸는 미래_스마트시티를 중심으로_1부
학습자료	<p>스마트시티 주목 이유</p> <p>1. 인구 · 경제의 변화다. → 세계적 저출산 · 고령화 현상 심화 _ 인구구조 변화 야기 → 저성장 장기화 예상 _ 대규모 도시개발사업 효용성 저하 예상</p> <p>2. 기후 · 환경의 변화 → 도시가 온실가스 배출 주범 _ 친환경 도시 국제적 관심 → <u>지구 전체 면적 도시 비중 2%, 지구 온실가스 70% 도시 생산</u></p> <p>3. 기술 발전 → 디지털 기술 기반 4차 산업혁명 _ 효율적 스마트시티 구현 → 4차 산업혁명 핵심 디지털 기술 _ 도시 내 초연결 · 초지능 가시화 → 다양한 형태 데이터 공유 _ 지능형 platform 경제 핵심</p>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
7 차시	빅데이터가 꿈꾸는 미래_스마트시티를 중심으로_2부
학습자료	<p>Smart City _ 일상의 변화 / 스마트 교통 → 2012년 '빅데이터 R&D 이니셔티브' 발족</p> <ul style="list-style-type: none">→ 도로 환경센서 데이터 수집 _ 안전운전 보조 서비스 등장→ 교통 및 주차난 _ Mobility as a Service 플랫폼 활용<ul style="list-style-type: none">_ 유희 주차공간 정보 공유 _ 주차난 해소



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
7 차시	빅데이터가 꿈꾸는 미래_스마트시티를 중심으로_2부
학습자료	<p>Smart City _ 일상의 변화 / 에너지 자립화</p> <p>→ 재생에너지 인프라 확충 _ 도시 내 필요한 전력 스스로 생산 _ 스마트그리드 이용, 전력 거래 가능 추진</p> <p>→ 도로태양광 _ 프랑스, 스웨덴, 네덜란드, 중국 등 시범사업 추진 _ 세종시 자전거도로 태양광 지붕 세계 우수사례</p>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
7 차시	빅데이터가 꿈꾸는 미래_스마트시티를 중심으로_2부
학습자료	<p>우리나라 스마트 시티 구축 필요성</p> <p>1. 인구 및 경제 분야</p> <ul style="list-style-type: none">→ 저출산, 고령화, 저성장, 공유경제 등 산업구조 변화_ 대규모 개발수요 감소 및 도시의 쇠퇴, 공간의 격차 발생 예상_ 신규투자 감소, 인프라 노후화, 공유경제 등 패러다임 촉진 전망 <p>2. 기후 및 환경 분야</p> <ul style="list-style-type: none">→ 지속 가능 도시 모델_기후변화로 재난·재해 위험성 점증_ 미세먼지 등 환경문제는 국민 안전을 크게 위협하는 요소 작용 <p>3. 기술 및 산업 분야</p> <ul style="list-style-type: none">→ 초연결·지능사회 출현, 신산업 대두_도시 내 초연결, 초지능화 가속→ 다양한 신산업 출현, 산업구조 혁신, 일자리 수요 변화_ 스마트시티 구축 정책 필연



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
7 차시	빅데이터가 꿈꾸는 미래_스마트시티를 중심으로_2부
학습자료	<p>제3차 스마트도시 종합계획 / 2019~2023</p> <p>1. 세종시 사례</p> <ul style="list-style-type: none">→ 공공데이터 기반 7대 서비스 구현 최적화 공간 계획 마련_ 각 병원 네트워크 구축하여 환자에게 최적화된 병원 안내_ 응급 상황의 경우 골든 타임 확보 도움_ 온라인 교육, 로봇 관련 교육, 맞춤형 쇼핑 등 다양한 서비스 제공 <p>2. 부산 에코델타시티 사례</p> <ul style="list-style-type: none">→ 로봇 및 물 관리 관련 신산업 육성 중점적 추진_ AI, 배송, 주차, 의료로봇 도입, 취약계층과 영세 상공인 적극 활용_ 스마트 물 관리기술 도입 _ 물 순환 최적화 실현



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
8 차시	정부는 빅데이터로 뭘 할까 _ 공공부문
학습자료	<p>기상데이터와 농산물 생산성 예측 모델 _ 기상청, 농촌진흥청</p> <ol style="list-style-type: none">1. 농수산물의 수매를 예측하기란 매우 어려우며, 그 이유는 기후와 시장 상황 변동으로 인하여 매년 달라지기 때문이다.2. 정부가 수급조절을 위해 노력하고 있으나 역부족으로, 농산물 가격을 예측할 수 있다면 수급조절이 가능하다.3. 기상청과 농촌진흥청은 농작물에 영향을 미치는 기상요인을 분석하여 '농작물 생산성 예측 모형'을 개발하였다.4. 이 모델은 <u>기상기후 데이터와 농산물 가격 데이터를 분석하여</u> <u>두 인자 간 상관관계 분석을 통해 상관성을 도출한 것이다.</u>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
9 차시	기업은 빅데이터로 뭘 할까 _ 마케팅부문
학습자료	<p>빅데이터 마케팅 _ Google Flu Trends</p> <ol style="list-style-type: none">1. 빅데이터 분석 및 활용 _ 미 질병관리본부보다 빠른 감기 예방 홍보2. 감기 증상 사람 증가 _ 감기 관련 단어 검색 빈도 증가 발견 → 독감 검색빈도 높은 지역 지도 표시 _ 독감 확산예측 가능 판단3. 감기 바이러스 확산 상황 _ 시간별, 지역별 지도 표시 → 독감 확산 _ 미국 보건당국 보다 1-2주 빠른 예측4. 플루 트렌드 서비스 제작으로 효과적 홍보 효과 → 혁신적 IT기업 이미지 강화5. <u>2013년 플루 대유행 시기 140퍼센트 오류 _ 구글 서비스 중단</u>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
10 차시	기업은 빅데이터로 뭘 할까 _ 제조부문
학습자료	<p>Industry 4.0</p> <p>독일 신제조업 전략</p> <p>→ 독일 Hannover Messe 2011에서 언급 _ 독일 신제조업 전략</p> <p>→ 메르켈 총리 _ 단순하고 상징적 표현인 Industry 4.0 결정</p>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
10 차시	기업은 빅데이터로 뭘 할까 _ 제조부문
학습자료	<p>독일이 인더스트리 4.0을 준비하게 된 계기</p> <ol style="list-style-type: none">1. 독일은 Made in Germany라는 완벽한 품질을 자랑하는 세계적인 제조업 강국임에도 불구하고,2. 첨단 ICT기술과 소프트웨어 역량을 보유한 미국과,3. 제조업분야에서 무섭게 급성장하고 있는 중국에 대해 큰 위기의식을 가졌기 때문이다.



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
10 차시	기업은 빅데이터로 뭘 할까 _ 제조부문
학습자료	<p>Smart Factory _ 센서</p> <ol style="list-style-type: none">1. 생산 과정 빅데이터 도입 _ 센서의 발달 의존2. 센서와 RFID 장착 _ 사용 기계의 모든 필수적 정보 관리자 전송3. 제품 생산 전 과정 모니터링 _ 자동 작업 지시, 생산 효율성 증가



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
10 차시	기업은 빅데이터로 뭘 할까_ 제조부문
학습자료	<p>제조업과 빅데이터 융복합 모델 _ Digital twin</p> <ol style="list-style-type: none">1. <u>가상물리생산시스템, 현실 설비를 디지털로 동일하게 구현</u>2. IoT 기술, 데이터 처리기술 발전 _ 세밀한 디지털 트윈 구현 가능3. 센서를 통해 설비 데이터 즉각 반영 _ 이상 징후 사전 감지 및 에너지 낭비 요인 제거 가능4. 생산 전 시뮬레이션 테스트 및 수정 _ 제품 생산 시 완성도 증가



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
10 차시	기업은 빅데이터로 뭘 할까_ 제조부문
학습자료	<p>세계 제조기업 빅데이터 도입</p> <ol style="list-style-type: none">1. 소비자 니즈에 적절한 최적의 생산 모델 구축2. Digital twin 도입 _ 시뮬레이션으로 문제점 개선, 출시 기간 단축 _ 다품종 소량 생산 가능3. <u>수율 개선 _ 생산성 향상 및 비용 절감</u>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
11 차시	나는 빅데이터로 뭘 할까? _ 의료부문
학습자료	<p>보건의료 빅데이터 _ 데이터 3법 개정</p> <ol style="list-style-type: none">1. 가명 정보 _ 개인의 이용 동의 없이 활용 가능2. 각 기관 분산 빅데이터 De-Identification _ 민간 연구자 제공 가능3. 맞춤형 의료서비스, 임상 의사결정 지원 _ 활용 목적 제공 가능4. 빅데이터 기반 제품, 서비스 질 개선, 신규 비즈니스 모델 발굴 가능



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
11 차시	나는 빅데이터로 뭘 할까? _ 의료부문
학습자료	<p>우리나라 보건의료 빅데이터 특징</p> <ol style="list-style-type: none">1. 전국민 단일 건강보험 체제, 세계 1위 보건의료 빅데이터 축적2. 의료기관 _ 전자의무기록 / EMR 도입, 디지털 의료정보 보유3. <u>공공과 민간 보건의료 빅데이터 _ 활용 초기 단계</u>4. 소유권, 개인정보보호, 표준화 부족 등 많은 이슈들 산재



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
11 차시	나는 빅데이터로 뭘 할까? _ 의료부문
학습자료	<p>보건의료데이터 현황</p> <ol style="list-style-type: none">1. 건강보험공단 _ 보험료, 진료, 검진 등 약 3조4천억 건 데이터 보유2. 건강보험심사평가원 _ 진료, 투약, 의약품 등 약 3조건 축적3. 국립암센터 _ 국내 암 발생 현황 일괄 보유4. <u>민간 의료기관 90% EMR 구축 _ 막대한 진료정보 보유</u>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
11 차시	나는 빅데이터로 뭘 할까? _ 의료부문
학습자료	<p>보건의료 빅데이터 개방과 활용 역기능</p> <ol style="list-style-type: none">1. 유전정보 자체가 개인식별정보 _ 연구 대상으로 암호화 불가능2. 민간 기업 보유 다른 개인정보 결합 _ 개인 식별 가능3. 보건의료데이터로 얻은 기술·경제적 성과 _ 민간 기업 독점



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
11 차시	나는 빅데이터로 뭘 할까? _ 의료부문
학습자료	보건의료 빅데이터 활용 <ol style="list-style-type: none">1. 활용목적 외 다른 목적 활용 금지2. 허가받지 않은 데이터의 유통이나 거래 금지3. 관련 법규 준수 및 기업윤리 실천



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
12 차시	기업은 빅데이터로 뭘 할까_통신, 에너지, 금융, 항공, 교육부문
학습자료	빅데이터 활용 _통신부문 1. 데이터 과학 적용 _ 효과적 마케팅, 사업전략 수립 이익 극대화 2. 데이터 송수신 _ 그 양이 폭발적으로 증대 3. <u>모바일 네트워크 _ 현재 전 세계 인구 95% 연결</u>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
12 차시	기업은 빅데이터로 뭘 할까_통신, 에너지, 금융, 항공, 교육부문
학습자료	<p>빅데이터 활용 _ 통신부문 / 비식별화 이용자 단말기 위치정보</p> <ol style="list-style-type: none">1. COVID-19 확산방지 정책 및 전염병 확산 사례 연구 활용2. 이동제한 정책 시행 시 정책효과 여부 확인 활용3. <u>세계 각국, 전염병 확산 연구 모델링 연구 활용</u>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
12 차시	기업은 빅데이터로 뭘 할까_통신, 에너지, 금융, 항공, 교육부문
학습자료	<p>빅데이터 활용 _ 통신부문 / 서울시 생활 이동 데이터</p> <ol style="list-style-type: none">1. <u>서울 내 모든 사람들의 이동 집계 데이터</u>2. 통근시간과 통학시간 개선 개발, 서울시 각종 정책 수립 활용 예정3. 버스노선 증차 결정 및 버스노선 신설4. 청년 공공주택 부지 선정 _ 도시 공간 구조개선 활용 전망



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
12 차시	기업은 빅데이터로 뭘 할까 _ 통신, 에너지, 금융, 항공, 교육부문
학습자료	<p>빅데이터 활용 _ 에너지부문 / 전력 빅데이터 활용</p> <ol style="list-style-type: none">1. 전봇대에 스마트 센서 부착, 전력 데이터 실시간 수집2. 폭염 전력 이용량 예측 _ 단계별 에너지 절감 문자메시지 전송3. 독거노인 전력 사용량 감소 _ 사회복지사에게 자동으로 통보4. 지진 발생 시 충격 감지정보 수집 _ 사고 대비 · 복구 활용5. 미세먼지 빅데이터 수집 _ 정부부처, 지방자치단체 서버 전송



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
12 차시	기업은 빅데이터로 뭘 할까 _ 통신, 에너지, 금융, 항공, 교육부문
학습자료	<p>빅데이터 활용 _ 금융부문</p> <p>1. 신용정보원 '크레디비(CreDB)'</p> <ul style="list-style-type: none">→ 금융권 빅데이터 공유 서비스 _ 무료 공개→ 개인정보가 유출되지 않도록 데이터를 재가공해 공개→ 핀테크 스타트업 고객 맞춤형 서비스개발, 보험사는 보험 상품 구성 <p>2. 신용카드사</p> <ul style="list-style-type: none">→ 빅데이터 활용 마케팅 _ 충분한 빅데이터 보유 신용카드 회사→ 신한카드 '초개인화 서비스' _ Time, Place, Occasion 예측_ 알고리즘과 플랫폼 기반 맞춤형 혜택 제공 서비스 <p>3. 시중 주요 은행</p> <ul style="list-style-type: none">→ <u>기업 여신 심사 빅데이터 활용 _ 중요 정보 분석 및 부실 징후 예측</u>→ <u>빅데이터와 AI 이용 _ 빠르고 정확한 분석 및 체계적인 리스크 관리</u> <p>4. 보험사 _ 콜 센터</p> <ul style="list-style-type: none">→ 고객 불만 또는 문의 _ 선제적 대응 시스템 구축, 보험사기 예측 및 적발→ 빅데이터 기반 목소리 자동인식기법_ 통화



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
13 차시	나는 빅데이터로 뭘 할까 _ 창업
학습자료	<p>빅데이터 활용 창업 _ 공공데이터 포털</p> <ol style="list-style-type: none">1. 공공데이터 검색, 추천 서비스 도입 → 2021년 4월 공공데이터 포털 - 추천 데이터 서비스 시행 → '공공데이터 큐레이션 서비스' _ 2021년 5월 한시적 실시2. 공공기관의 공적인 정보, 무료이며 누구나 이용 가능 → 공공데이터포털 _ 국민에게 개방할 공공데이터가 모두 있는 공간3. 창업 비즈니스모델 활용 공공데이터 검색 가능 → 공공데이터 목록 _ 데이터 명칭, 키워드, 설명 등 포함 검색 → 필요한 데이터 유형별, 기관별 자유롭게 검색 가능



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
13 차시	나는 빅데이터로 뭘 할까 _ 창업
학습자료	<p>우리나라 보건의료 빅데이터 특징</p> <ol style="list-style-type: none">1. <u>공공데이터 신청이 반려된 경우 명확한 근거를 바탕으로 신청인에게 반려된 이유를 통보해 주며, 제공신청이 반려됐다면 분쟁조정 신청으로 한 번 더 공공데이터 요청이 가능하다.</u>2. 공공데이터 포털에서는 누구나 사용할 수 있도록 공개된 오픈 API를 제공하고 있다.3. 공공데이터 포털에서 오픈 API를 활용할 경우 개발계정 신청과 운영계정을 신청해야 한다.4. 공공데이터 시각화는 공공데이터를 제공받아 직접 제작하고 싶다면 공공데이터 시각화를 활용할 수 있다.



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
14 차시	누구나 알 수 있는 빅데이터 분석기법
학습자료	<p>데이터 마이닝 세부 분석 방법 연관성 분석과 분류(트리) 분석, 군집 분석, 예측 분석의 네 가지로 구분</p> <ol style="list-style-type: none">1. 연관성 분석 데이터 간에 어느 정도의 유사성이 있는지를 살펴보는 것2. 분류 트리 분석 데이터를 특정 기준으로 나누어 분류하는 것을 의미3. 군집 분석 특정 기준을 미리 설정 하지 않는 분석 기법4. 예측 분석 기존의 데이터 패턴을 기반으로 미래에 일어날 특정 사건을 예측하는 분석 기법



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
14 차시	누구나 알 수 있는 빅데이터 분석기법
학습자료	<p>빅데이터 분석</p> <ol style="list-style-type: none">1. 숨겨진 패턴과 알려지지 않은 정보를 찾는 과정2. 의사 결정 시 사용 _ 최선의 대안 선택, 근거 제시 중요 역할3. <u>파급 효과가 큰 의사 결정 _ 데이터 분석 바탕 의사 결정 적절</u>4. 효과적 데이터 분석 _ 빅데이터 분석 플랫폼 구축



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
15 차시	데이터가 스스로 답한다 _ AI
학습자료	<p>인공지능과 빅데이터</p> <ol style="list-style-type: none">1. 빅데이터가 없는 AI는 무용지물입니다. Why AI Would Be Nothing Without Big Data2. 데이터의 폭발적 증가로 엄청난 속도로 발전3. 매년 데이터 두 배 증가, 10년 후 1500억개 네트워크 센서 예측 → AI의 생각과 느낌 이해 도움, 학습 속도 향상, 데이터 분석 자동화4. 과거 AI _ 제한된 데이터 집합, 대량 데이터 몇 초내 분석 불가능 → <u>현재 _ 실시간 접근 가능한 데이터와 분석도구, 빠른 분석 가능</u>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
15 차시	데이터가 스스로 답한다 _ AI
학습자료	<p>인공지능과 빅데이터 양면성 _ 역기능</p> <ol style="list-style-type: none">1. Big Data와 AI 기술 _ 인간의 일자리 및 직업 활동 감소 위험2. <u>개인정보보호와 활용 _ 범죄에 활용되지 않도록 법률 정비 필요</u>3. 개개인에 대한 사생활 침해는 매우 우려4. 개인, 기업이 정보 독점 _ 사회 통제 권력 가능



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
16 차시	맛만 보는 빅데이터 시각화 기술
학습자료	<p>데이터 활용 능력</p> <ol style="list-style-type: none">1. <u>데이터 활용 능력은 데이터 분석과 데이터 시각화라고 볼 수 있다.</u>2. 데이터 분석은 기술적으로 데이터를 수집하고 정제하는 데이터 가공 단계부터 분석 기법을 활용한 데이터 분석에 이르는 과정에 필요한 능력을 말한다.3. 데이터 시각화는 분석 결과를 시각적으로 표현하여 분석 결과와 가치를 서비스에 어떻게 적용할 것인지, 또 그 가치를 사용자에게 어떻게 보여줄 것인가를 나타낸다.



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
16 차시	맛만 보는 빅데이터 시각화 기술
학습자료	<p>통계자료의 시각화 발전</p> <ol style="list-style-type: none">1. <u>17세기 지도와 그래프, 1800년대 초 파이도표 발명</u>2. 샤를 미나르 _ 나폴레옹의 러시아 침공 지도 제작 _ 군대 규모, 퇴각 경로, 당시 기온과 침략기간 서술3. 컴퓨터 _ 많은 양 데이터, 빠른 속도로 데이터 시각화 처리 → 과학과 예술의 조합으로 빠르게 진화, 기업 환경 획기적 변화



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
17 차시	필요한 데이터는 나에게 맡겨라_IoT
학습자료	<p>데이터 시각화의 장점</p> <ol style="list-style-type: none">1. 시각적 표현 analytics 정보 바탕, 개념 이해, 패턴 식별2. 대화형 시각화 기술_차트와 그래프 검토, 개념 심층적 이해 가능3. 분석 결과, 적절한 차트나 이미지 등 시각화 <p>→ <u>정보를 빠르게 확산, 자료에 대한 기억 증가, 효과적 커뮤니케이션</u></p>

훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
17 차시	필요한 데이터는 나에게 맡겨라_IoT
학습자료	<p>데이터 시각화 _ 기능</p> <ol style="list-style-type: none"> 1. 정보를 빠르게 이해하는 기능 → 명확, 일관적 방식으로 많은 양의 데이터를 파악하여 결론 도출 → 스프레드시트 분석에 비해 빠른 분석 가능, 질문에 대한 해답 도출 2. 관계성과 패턴 파악 기능 → 방대, 복잡한 데이터 _ 그래픽 표현 → 쉬운 이해, 관련성 높은 매개 변수 인지, 목표에 영향을 주는 영역 집중 3. 새로운 동향 파악 기능 → 비즈니스와 시장 동향 파악 _ 경쟁 우위 확보 및 수익 증가 → 제품 품질, 고객 이탈 영향 이상치 발견 _ 문제 확대 전 사전 해결 4. 정보 전달 기능 → 새로운 분석 정보 발견 _ 해당분야 사람들과 공유 → 도표나 그래프 등 시각적 데이터 형식 사용 _ 신속한 메시지 전달



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
17 차시	필요한 데이터는 나에게 맡겨라_IoT
학습자료	<p>데이터 시각화 필요성</p> <ol style="list-style-type: none">1. 많은 양의 데이터를 한 눈에 파악 → 공공부문, 금융, 의료 등 _ 데이터로만 현상 파악, 예측 곤란 → 반대한 데이터 _ 시각적 요소 활용, 요약 표현 한 눈에 파악 가능2. 누구나 쉽게 데이터 인사이트 → 인간은 획득하는 정보의 80% 이상이 시각을 통해 취득 → 시각화를 통한 데이터 활용 범위 및 새로운 가치창출 범위 확대 → 시각화를 통한 스토리텔링화 _ 사람에게 전달 시 가장 효과적3. 정확한 데이터 분석 결과 도출 → 정확한 분석을 위한 데이터 탐색 방법 활용 → 요약 통계 정보 _ 시각화를 통한 정확한 분석 가능4. 데이터 기반 의사결정 → 빠른 인지, 기억, 메시지의 스토리텔링에 대한 공감 _ 분석가뿐만 아니라 비전문가도 데이터 활용 가능 _ 직관이 아닌 데이터를 근거로 중요한 Datadriven decision 가능5. 활용할 수 있는 많은 분야 → 데이터 시각화를 데이터 활용 방법으로 사용 → 데이터 개방포털 시각화서비스 _ 개인데이터 활용 시각화 차트 제작



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
17 차시	필요한 데이터는 나에게 맡겨라_IoT
학습자료	<p>빅데이터와 사물인터넷</p> <ol style="list-style-type: none">1. 빅데이터는 IoT 생태계의 핵심이라고 할 수 있다.2. IoT 기술이 추구하는 생활이란, 사물들 간의 통신을 통해서 이전에는 인간이 스스로 행동했던 것들을 기계가 스스로 작동하게 만들어 생활을 편리하게 만드는 것을 말한다.3. 수년 내 엄청난 데이터 생성 예상 → 대량 정보가 공유되고 융합될 수 있는 접근 필요 “폭발적 정보 증가는 필연적으로 빅데이터 기술 결합이 요구”4. IoT와 빅데이터는 결국 하나의 기류로 합쳐질 것으로 예상 “수년 내 이루어질 IoT와 빅데이터의 융합은 우리 생활에 큰 영향”



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
17 차시	필요한 데이터는 나에게 맡겨라 IoT
학습자료	<p>IoT의 정의</p> <ol style="list-style-type: none">1. IoT는 Internet of Things의 줄임말로 직역하면 사물들의 인터넷을 말하며,2. 먼 미래에는 정말 모든 사물들에 인터넷 연결을 적용하고자 하는 것이 최종 목적이 되고,3. 사물들에 인터넷이 연결된다는 것은 연결된 기기끼리 통신이 가능해진다고 할 수 있다.



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
17 차시	필요한 데이터는 나에게 맡겨라_IoT
학습자료	<p>사물인터넷과 빅데이터 융복합 형태 사례</p> <p>1. 코마츠 사례 → 중앙관리센터 _ 모든 장비 센서 수집, 관련 정보 고객 제공 서비스 → 실시간 데이터 수집 및 분석 _ 고객과 대리점, 정보 전송</p> <p>2. 맥키니 사례 → 빌딩 내 수만 개 센서 설치, 데이터 수집 및 분석, 실시간 모니터링 _ 비효율적 에너지 사용 문제점 발견 및 분석을 통한 해결 _ 전체 에너지 사용의 20% 정도 절감 효과</p> <p>3. 보건의료분야 / 미국 의료기관 사례 → 블루투스 청진기 및 자율주행 카트 등 _ 사물인터넷 기술 적용 → 수액 모니터링 시스템 _ 환자별 투약 상태 확인 _ 간호사, 남은 수액량 파악 시스템에서 확인 가능 → 격리환자 모니터링 _ 조명 자동 on/off, 환자상태 실시간 그래프 제공 _ 의료진 외부 모니터링, 홈 케어 서비스 적용 예정</p>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
18 차시	데이터 저장의 새로운 패러다임
학습자료	<p>데이터 레이크 주목 배경</p> <ol style="list-style-type: none">1. 스마트 폰과 SNS, IoT에서 다양한 데이터가 방대하게 생성되므로 이를 간단하고 효율적으로 관리할 플랫폼이 필요하며, 데이터 레이크는 그 수단이 될 수 있다.2. 기업이 다양한 데이터 소스를 토대로 의사 결정을 내리는 데 도움이 된다.3. 조직의 시너지를 강화할 수 있도록 부서별 또는 분야별로 흩어진 데이터를 통합할 수 있다.4. 데이터웨어하우스보다 운영비가 덜 들어 경제적이다.5. 큰 비용을 쓰지 않고 쿼리 속도를 높일 수 있다.



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
18 차시	데이터 저장의 새로운 패러다임, 데이터 레이크
학습자료	<p>데이터 웨어하우스</p> <ol style="list-style-type: none">1. 미리 결정된 목적 _ 수집 및 생성 데이터 저장소2. 데이터 저장 전 미리 정의된 스키마 적용, 데이터 정리 및 구성3. 데이터 전 처리 _ 높은 수준 분석 용이, <u>비전문가 쉽게 사용</u>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
18 차시	데이터 저장의 새로운 패러다임, 데이터 레이크
학습자료	<p>데이터 레이크</p> <ol style="list-style-type: none">1. 원시 데이터 저장 저장소, 다양한 구조 데이터 저장2. 저장 데이터 요소 _ 고유한 식별자와 메타데이터 태그3. 필요 시 쉽게 쿼리, 미리 정의된 스키마 없음4 <u>데이터 과학자 등 사용, 고급 필터 및 분석 적용</u>5. 대규모에도 불구하고 상용 하드웨어 사용, 비용 저렴



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
19 차시	마지막으로 빅데이터 분석가가 되려면
학습자료	<p>빅데이터 분야 직업</p> <p>데이터 엔지니어, 데이터 분석가, 데이터 과학자 → 구분한 3가지 직무가 100% 다른 업무 진행이라고 단정하기 곤란</p>



훈련과정	빅데이터는 나의 친구 _ 나도 빅데이터 전문가
19 차시	마지막으로 빅데이터 분석가가 되려면
학습자료	<p>데이터 엔지니어, 데이터 분석가, 데이터 과학자의 직무 역량</p> <p>1. 데이터 엔지니어</p> <ul style="list-style-type: none">→ 데이터 관리 담당 _ 데이터 수집과 저장, 유지, 보수 등 업무 포함→ 데이터 활용을 위해 온전한 데이터셋을 빠르게 전달하는 목표→ 하둡, 맵리듀스 등 _ 데이터 시스템 이해 필요→ 관계형 데이터베이스, SQL _ 데이터 저장과 추출 방식 이해→ 데이터 활용을 위해 가장 중요하면서도 먼저 구축해야 하는 분야 <p>2. 데이터 분석가</p> <ul style="list-style-type: none">→ 데이터 활용 _ 조직이 목표를 달성하기 위한 서술적 분석→ 특별한 패턴 발견 _ 그 결과를 비즈니스적 활용→ 패턴 분석 _ 통계학적 지식 필요→ 데이터 분석 _ Python이나 R 등 코딩 능력 필요→ 가공 데이터 표현 _ 시각화능력 매우 중요 <p>3. 데이터 과학자</p> <ul style="list-style-type: none">→ 데이터를 통해 가치 발굴 목적 _ 분석 방식이 데이터 분석가와 차이→ 머신 러닝, 딥 러닝 기술 활용 _ 미래 예측 직무 수행→ 기본적 통계 분야 역량과 코딩 역량 구비→ 가장 중요한 역량 _ 머신 러닝, 딥 러닝 생태계 및 알고리즘 이해→ 인공지능 전문가가 가지고 있는 역량을 어느 정도 구비