

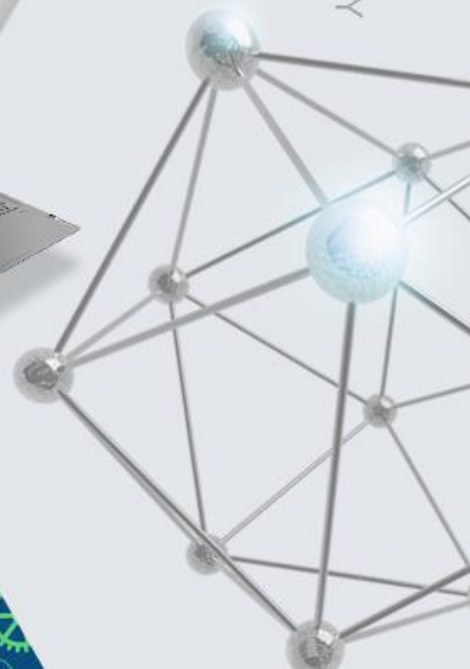


한국기술교육대학교  
온라인평생교육원

The 4th Industrial Revolution is characterized by super connectivity and super intelligence, where various products and services are connected to the network, and artificial intelligence and information communication technologies are used in 3D printing, unmanned transportation, robotics. Of the world's most advanced technologies.

# 빅데이터 입문

## 빅데이터 기술활용과 비즈니스 적용



The 4th Industrial Revolution is characterized by super connectivity and super intelligence, where various products and services are connected to the network, and artificial intelligence and information communication technologies are used in 3D printing, unmanned transportation, robotics. Of the world's most advanced technologies.



# 빅데이터 기술활용과 비즈니스 적용



## 학/습/목/표

1. 빅데이터 분석을 위한 활용 기술을 설명할 수 있다.
2. 비즈니스 적용에 빅데이터 분석방법을 활용할 수 있다.
3. 빅데이터 활용을 위한 이슈를 설명할 수 있다.



## 학/습/내/용

1. 빅데이터 분석을 위한 활용 기술 정의하기
2. 빅데이터 기술의 기업 활용하기
3. 빅데이터 활용을 위한 세 가지 이슈



### 1. 빅데이터분석을위한 활용 기술 정의하기

#### 1) 빅데이터 분석과 시각화

##### (1) 시각화 기술

- 데이터 분석 결과를 쉽게 이해할 수 있도록 시각적으로 표현하고 전달되는 과정

##### (2) 활용목적

- 이해관계자들에게 새로운 인사이트를 도출함
- 정보를 보다 명확하고 이해하기 쉽게 전달함

##### (3) 주요특징

- 데이터에 대한 직관적 이해 가능
  - 지각적 추론 가능
  - 데이터에 대한 신속한 접근 가능
- 일반 초보 사용자들도 쉽게 데이터 분석에 접근하도록 지원



### 1. 빅데이터분석을위한 활용 기술 정의하기

#### 1) 빅데이터 분석과 시각화

##### (4) 시각화의 원칙

- 데이터 분석 목적과 부합해야 함
- 다양한 정보를 전달할 수 있어야 함
- 사용자들이 원하는 정보를 전달해야 함

##### (5) 시각화의 수행 3단계

- 정보의 조직화: 사용자 정보인지에 관여하여 데이터를 분류하고 조직화, 체계화함
- 정보의 시각화: 사용자의 정보 지각에 관여하고 시각화를 통해 효율적인 정보전달을 모색함
- 사용자와 경험: 정보와 사용자의 상호작용을 통해 사용자 경험, 인사이트 도출과정에 지원함



### 1. 빅데이터분석을위한 활용 기술 정의하기

#### 1) 빅데이터 분석과 시각화

##### (6) 시각화의 방법

- 차트
- 프로그램
- 지도
- 일러스트레이션

##### (7) 데이터의 시각화 기술

- 시간의 시각화: 누적막대 그래프, 점 그래프
- 분포의 시각화: 파이 차트, 도넛 차트
- 관계의 시각화: 스캐터 플롯
- 비교의 시각화: 히트맵



### 1. 빅데이터분석을위한 활용 기술 정의하기

#### 2) 빅데이터 분석 도구

##### (1) SPSS

- 사용자 그래픽 인터페이스가 우수함
- 통계적 데이터 분석, 데이터 마이닝
- 비교적 소용량 데이터 처리에 적합함

##### (2) SAS

- 프로시저 기반 스크립트 가능
- 모듈을 결합하여 고급분석 수행



### 1. 빅데이터분석을위한 활용 기술 정의하기

#### 2) 빅데이터 분석 도구

##### (3) R(통계 인력들이 선호)

- 다양한 통계기법에 활용 가능, 다양한 시각화 도구 제공
- 빠른 처리 속도
- 무료 설치 (저렴한 비용)
  - R 프로젝트 웹 사이트(<https://www.r-project.org/>)

##### (4) PyThon(IT 개발자들이 선호)

- 간결한 문법, 인간의 사고체계와 유사한 표현구조
- 유지보수와 관리 용이
- 풍부한 라이브러리/함수
- 단점
  - 모바일 웹 환경에 적용이 어려움
  - 속도가 느림



## 2. 빅데이터 기술의 기업 활용하기

### 1) 시각화 기술의 활용

#### (1) 자동차 대시보드

- 자동차의 속도, RPM, 연료량, 오일, 타이어 공기압 체크
- 자동차 운행과 안전에 있어서 핵심적인 정보들을 실시간으로 시각화된 형태로 확인할 수 있음
- 빅데이터 분석을 통해 수 많은 데이터를 활용할 수 있도록 시각화하여 제시함

#### (2) 제조업 공장의 대시보드

- 공장의 관리자나 매니저가 그래프나 차트로 일목요연하게 파악 가능함
- 향후 취할 조치에 대한 판단의 근거로 활용함

#### (3) 인사 부서의 대시보드

- 근태기록의 문제와 불만 예측
- 인적자원이 이탈하여 퇴직하는 일을 사전에 방지 가능함





## 2. 빅데이터 기술의 기업 활용하기

### 1) 시각화 기술의 활용

#### (4) 마케팅 부서의 대시보드

- 광고비 증가? or 감소할 것인지 판단
- 인터넷, 유튜브, 네이버 등 채널 활용 여부 판단

#### (5) 시각화 기술의 활용

- 즉시 활용 가능하고 상대적으로 용이함
- 수집 데이터를 활용 가능하도록 정제함
- 기업의 상황을 분석하여 업무 도입의 동기부여함



## 2. 빅데이터 기술의 기업 활용하기

### 2) 연관분석 기법의 활용

#### (1) 연관분석

- 인과관계는 없으나, 공동으로 일어나는 동시 발생 이벤트나 행위를 파악하는 것

#### (2) 연관분석의 사례 - 공정관리

- 두 공정이 동시 일어나도록 작업 배치
- 근무방식 변경의 판단근거로 활용

#### (3) 연관분석의 사례 - 고객관리

- 별점 2개 이하의 사람들만 찾아 관리
- 악플 행위 등 비난 행위에 대한 선제적 대응

#### (4) 연관분석의 사례 - 인사 관리

- 결근이 잦은 사람은 집중 관리함
- 불만에 대한 선제적 대응으로 이탈행위 방지



## 2. 빅데이터 기술의 기업 활용하기

### 3) 군집화 기법의 활용

#### (1) 클러스터링

- 어떤 대상들을 그룹화하는 것

선거	유권 지역별로 군집화하여 집단에 적합한 선거 메시지 준비
마케팅	고객들을 분류하여 각 집단에 맞는 광고 메시지 개발
대학교의 학생관리	학생들을 군집화하여 집단에 맞는 수업방법 적용
기업의 인사관리	직원들을 군집화하여 적합한 교육, 근무 방식 등 인사관리 적용

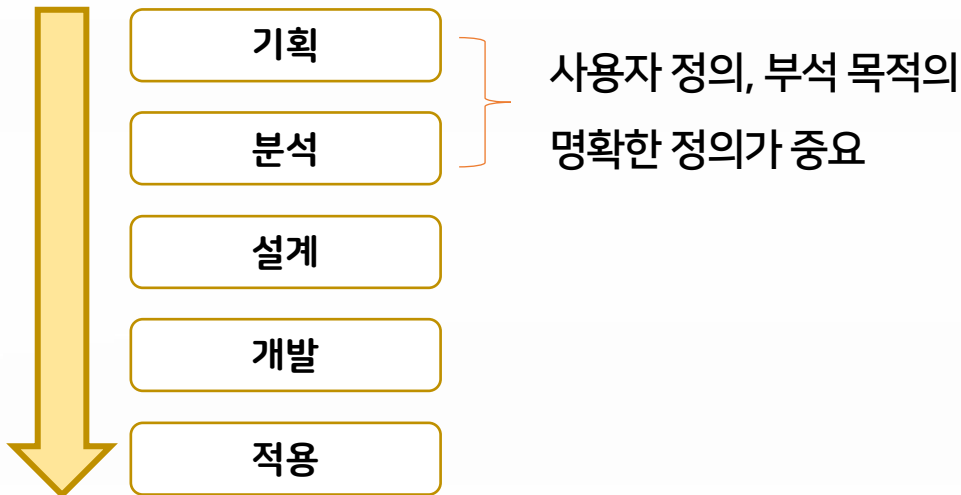


### 3. 빅데이터 활용을 위한 세 가지 이슈

#### 1) 사용자 정의와 명확한 분석목표의 정의

##### (1) 사용자/분석 목표의 정의

- 사용자가 누구인가?
- 그들이 원하는 것이 무엇인가?



- 빅데이터 분석 결과물에 대해 사용자인 현업, 실무자들과 분석을 기술적인 측면에서 지원
- 서로의 입장을 이해하고 공동으로 작업하는 것이 중요



### 3. 빅데이터 활용을 위한 세 가지 이슈

#### 2) 비즈니스 모델 측면의 접근

##### (1) 비즈니스 모델 측면의 접근 중요성

- 빅데이터 기반 비즈니스의 3가지 구성요소

- ① 빅데이터 기술: 가트너 그룹의 네 가지 분류
- ② 비즈니스 모델: 고객, 제공할 가치, 수익
- ③ 빅데이터 마케팅: 빅데이터 기술을 어떻게 차별화하여 쓸 것인가?

→ 세 가지에 대한 균형 필요

- 우리의 고객, 제공할 가치, 수익 창출 방법에 대한 비즈니스 모델 측면의 명확한 정의 필요



### 3. 빅데이터 활용을 위한 세 가지 이슈

#### 3) 데이터 공유와 활용

##### (1) 데이터 공유와 활용

- 일반적으로 데이터가 없어서가 아닌, 데이터가 공유되지 못하는 것이 문제
- 공유하고자 하는 의지/마인드 발생
- 최고경영자의 관심, 지원과 정책 필요

##### (2) 데이터 3법 개정 (2020년)

- 개인정보보호법/정보통신법에 대한 시대적 트렌드에 부합하는 개정과 적용 필요
- 4차 산업 혁명 도래에 맞춰 개인과 기업이 정보를 활용하는 폭을 넓히기 위해 마련
  - ① 개인정보보호법
  - ② 정보통신망법
  - ③ 신용정보법



### 3. 빅데이터 활용을 위한 세 가지 이슈

#### 3) 데이터 공유와 활용

##### (3) 공공데이터

- 우리나라 공공기관이 업무를 수행하면서 만들어낸 다양한 데이터
- <https://www.data.go.kr/>에서 확인 가능
- 마스크 재고 앱, 학교 정보 제공, 아이 놀이 정보 제공 등  
공공 데이터를 활용하여 새로운 가치를 만들어냄

##### (4) 마이 데이터

- 개인이 자신의 정보를 적극적으로 관리, 통제하는 것은 물론 이러한 정보를 신용이나 자산 관리 등에 능동적으로 활용하는 일련의 과정을 의미
- 자신의 정보를 한꺼번에 확인하고, 나에게 맞는 맞춤 상품이나 서비스를 추천 받음
- 나 자신(정보 주체의 중심)을 데이터의 주인으로 전환하는 것
- 데이터를 활용하여 새로운 가치를 생성하고 사용자에게 제공



### 3. 빅데이터 활용을 위한 세 가지 이슈

#### 3) 데이터 공유와 활용

##### (5) 마이 데이터 활용 사례

아파트 구입	<ul style="list-style-type: none"><li>어느 지역의 어떤 아파트를 구매?</li><li>대출은 얼마를 받으면 좋은가?</li><li>어느 기관에서 받아야 유리한가?</li></ul>
건강 정보	<ul style="list-style-type: none"><li>주의 깊게 관리해야 될 질병은?</li><li>어느 의료기관 병원에 가는 게 좋은가?</li></ul>
취업정보	<ul style="list-style-type: none"><li>어떤 자격증을 취득하면 좋은가?</li><li>취득하기 위한 비용은 얼마인가?</li><li>어디에 취업할 수 있는가?</li></ul>





### 1. 빅데이터 분석을 위한 활용 기술 정의하기

#### 1) 빅데이터 분석과 시각화

- 정의 : 데이터 분석 결과를 쉽게 이해할 수 있도록 시각적으로 표현하고 전달되는 과정
- 활용목적 : 그래프와 차트를 통하여 정보를 명확하고 이해하기 쉽게 전달하고 새로운 인사이트 도출

#### 2) 빅데이터 분석 도구

- SPSS : 사용자 그래픽 인터페이스(User Graphic Interface)가 우수
- SAS : 모듈별 기능이 존재하고, 각각의 모듈을 결합하여 고급분석 수행 가능
- R : 다양한 통계기법에 활용이 가능하고, 다양한 시각화 도구 제공
- 파이썬 : 문법이 간결하고, 유지 보수와 관리가 용이



## 2. 빅데이터 기술의 기업 활용하기

### 1) 빅데이터 분석을 위한 활용기술

- 시각화 : 자동차 대시보드, 제조업 공장의 대시보드, 인사부서의 대시보드
- 연관분석 : 공정관리, 고객 관리, 인사 관리
- 군집화 : 선거, 마케팅, 대학교의 학생관리, 기업의 인사 관리





### 3. 빅데이터 활용을 위한 세 가지 이슈

#### 1) 빅데이터 분석 활용의 이슈

- 사용자/분석 목표의 정의 : 사용자가 누구이고 그들이 원하는 것이 무엇인지의 명확한 정의
- 비즈니스 모델 측면의 접근 중요성
- 데이터 공유와 활용