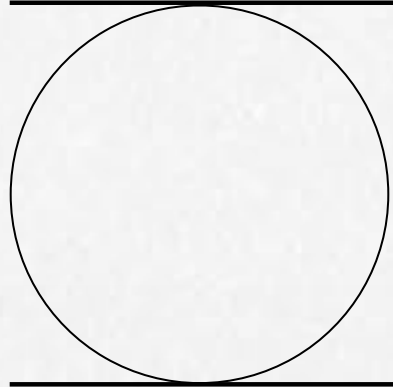
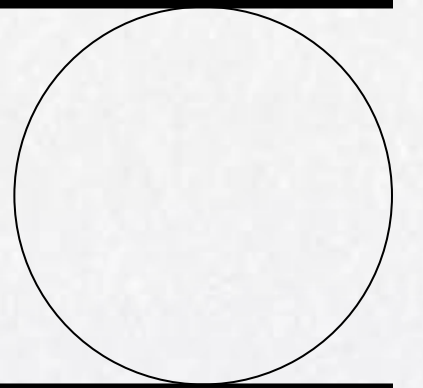
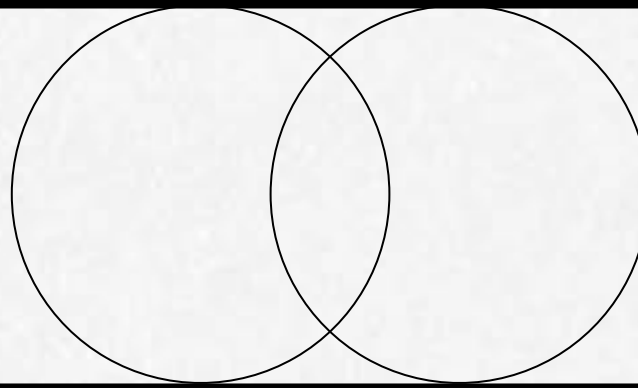


2023.05.25

---

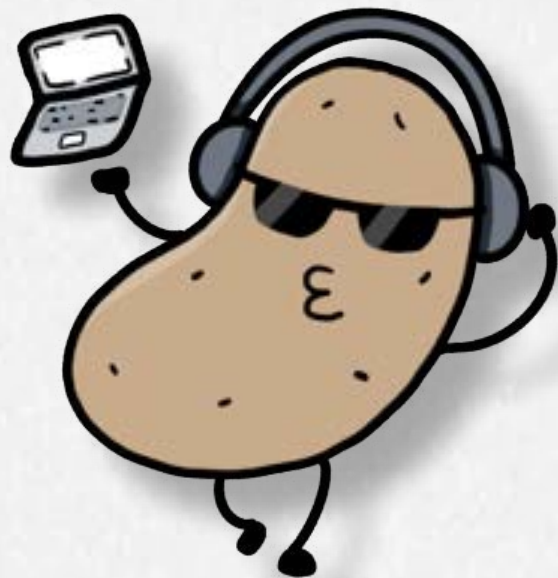
# 코딩공감자

---



# SELF INTRODUCTION

---



코딩공감자

---

# 코딩공감자의 진혹거 ✨

이 중에서 2개는 거짓입니다!  
진실일 것 같은걸 골라서 말해주시면 됩니다~

●  
●  
●

1

.

2

.

3

CDG GAMZA

코딩공감자는  
공대 소속 동아리  
이다.

우리는 모두 다  
21학번이다.

우리는 오늘 점심을  
다같이 모여서  
먹었다.

# 이수선생님의 진혹거 ✨

이 중에서 2개는 거짓입니다!  
진실일 것 같은걸 골라서 말해주시면 됩니다~

●  
●  
●

1

.

2

.

3

CDG GAMZA

이 중에서 가장  
나이가 어리다.

헤원여고  
졸업생이다.

공과대학에  
재학중이다.



# 이민채선생님의 진혹거 ✨

이 중에서 2개는 거짓입니다!  
진실일 것 같은걸 골라서 말해주시면 됩니다~

●  
●  
●

1

.

2

.

3

CDG GAMZA

20학번이다.

과를 전과했다.

7월에도 학교를  
다닐 예정이다.

# 정지혜 선생님의 진혹거 ✨

이 중에서 2개는 거짓입니다!  
진실일 것 같은걸 골라서 말해주시면 됩니다~

⋮

1

.

2

.

3

과학중점고등학교를  
졸업했다.

휴먼기계바이오  
공학부를 복수전공  
하고있다.

언니가 있고  
둘째이자  
막내이다.

# 박신영선생님의 진혹거<sup>☆</sup>

이 중에서 2개는 거짓입니다!  
진실일 것 같은걸 골라서 말해주시면 됩니다~

●  
●  
●

1

.

2

.

3

CDG GAMZA

위에 오빠만  
두명이 있다.

미팅에 성공한  
경험이 있다.

TV에 나온적이  
있다.



# MBTI ✨ 맞춰보기

이수

이민채

정지혜

박신영

# con

01 — 빅데이터란

02 — 엔트리란

# tents<sup>✦</sup>





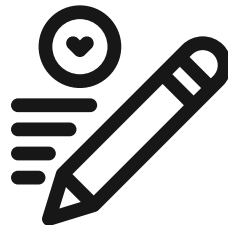
# BIGDATA 특징 ✨

이 곳에 알맞은 내용을 적어주세요.  
사용한 폰트는 G마켓 산스 Medium 16pt 입니다.



## 1 VOLUME

다양한 소스에서 방대한  
데이터를 수집  
- 1 tera~ 수십 peta



## 2 VELOCITY

매우 빠른 속도로 데이터  
가 유입되며 적시에 매우  
빠르게 처리 되어야 한다



## 3 VARIETY

다양한 유형의 형식으로 제공  
- 정형, 반정형, 비정형 데이터



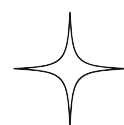
## 4 VALUE/ VISUALIZATION

회사 성장에 기여할 수  
있는 사람일까?

# 빅데이터 전문가 송길영이 말하는 '빅데이터란 무엇인가?'

<https://www.youtube.com/watch?v=2A8NelGuWgl>





# 빅데이터?

- 규모가 너무 크고 빠르거나 복잡하여 기존 방법으로는 처리하기 어렵거나 불가능한 데이터
- 이러한 데이터를 처리하는 관련 기술과 인력을 통칭하는 의미로 사용되기도 함

## Task

비즈니스 트랜잭션, 스마트 장치, 산업 장비, 비디오, 소셜 미디어 등 다양한 소스에서 방대한 데이터를 수집

## Behavior

사물 인터넷의 성장과 함께 데이터는 전례 없는 속도로 데이터가 유입되어 적시에 매우 빠르게 처리되어야 함

## Result

데이터는 기존 데이터베이스의 구조화된 숫자 데이터에서 구조화되지 않은 텍스트 문서, 이메일, 비디오, 오디오, 주식 시세 데이터 및 금융거래에 이르기까지 다양한 유형의 형식으로 제공



# 빅데이터는 어디에서 생성되는가?



01  
**비즈니스 데이터**  
기업에서 생성되는  
고객 및 구매와  
관련된 데이터

02  
**멀티미디어 데이터**

각종 멀티미디어  
콘텐츠로부터의  
이미지 및 비디오  
데이터



03  
**소셜미디어 데이터**  
트위터, 페이스북,  
블로그 등의  
소셜미디어 데이터

04  
**사물인터넷 데이터**

각종 센서에서 수집  
되는 M2M / 사물인  
터넷 스트림 데이터



05  
**의생명 데이터**  
의료정보, 생체정보,  
유전자 등의 데이터  
→ 바이오데이터

# 빅데이터 처리과정



## 수집

- 다양한 데이터 소스로부터 데이터를 추출하고 분석하기 용이한 형태로 저장

## 저장

- 빅데이터 저장을 위해서는 클라우드 시스템과 하둡이 활용
- 실시간으로 발생하는 데이터를 위해 인메모리 데이터 베이스 등 빠른 저장과 처리장치가 필요

## 데이터 정제 및 전처리

- 수집된 데이터는 바로 분석에 사용할 수 없는 경우가 대부분
- 단위의 차이, 결측값, 오류 데이터 등의 보정 필요
- 수집된 데이터를 분석이 가능한 형태로 정돈하는 과정

## 데이터 탐색

- 가벼운 데이터 분석, 요약 통계 (평균, 표준오차, 빈도, 최대값...)
- 전반적인 데이터의 내용을 파악하는 단계

## 분석

- 저장된 데이터에 대해 다양한 분석을 수행하는 과정
- 정형, 비정형의 대규모 실시간 데이터에 대한 분석 기법이 필요  
-> 빅데이터 분석기법

## 시각화

- 분석된 결과를 직관적으로 이해하고 의사 결정에 활용할 수 있도록 시각화가 필요

# 엔트리란?



1

.

2

.

3

네이버 커넥트재단에서 개발하고  
운영하는 비영리 소프트웨어  
교육 플랫폼

약 200만명 이상의  
유저들이 이용 중

2018년부터 2015 개정 교육과정에 따라  
엔트리가 교육용 코딩 언어로 채택되어,  
전국의 초등학생과 중학생들을 대상으로 한  
소프트웨어 교육을 시작함



# 엔트리 학습하기

## 이상한 티파티

●  
●  
● <https://playentry.org/learn?discoveryCategory=game>



# 엔트리 학습하기

## 여왕의 정원

●  
●  
● <https://playentry.org/learn?discoveryCategory=game>



코딩공감자

✧ We wish you  
✧ success. ✧

( 당신의 성공을 기원합니다. )