



AI 융합전문가 4차시 데이터분석 I

2024.09.29..





현재 하는 일



소속	주요 업무
(주)제타데이터 대표이사	데이터 분석, 전략 및 컨설팅, 데이터 가치평가 ODA 컨설팅 (Official Development Assistance)
(주)지구파트너스 감사	창업보육, 투자, 기업·기술가치평가, 사업타당성 분석
(주)메타로직 컨설팅 수석	ISP 컨설팅 (Information Strategy Planning) ISMP 컨설팅 (Information System Master Plan)

◎ 자격증

1. 경영지도사31기 (인적자원, 2016)
2. 창업보육매니저 (BI협회, 2018)
3. 기업·기술가치평가사 (기업·기술가치평가협회, 2018)
4. 기업재난관리사 실무과정 (행정안전부, 2019)
5. 데이터분석 준전문가 ADsP (데이터산업진흥원 K-Data, 2021)
6. 빅데이터 분석기사 (과학기술정보통신부 · 통계청, 2021)
7. 국제공인컨설턴트 CMC (ICMCI, 2023)
8. 인공지능(AI) 활용마스터1급 (뉴미디어교육연구소, 2024)



데이터 관련 비즈니스

한국경제

Hankyung.com
가업을 잇는 생각

알고 보면 재미있는 경영이야기 - 기업 경영과 자본

▲기업철 홍용기 전문위원

대중에서 이익을 창출하는 순서가 있어서도 경영이익에서, 타인자본에 대한 이자비용을 먼저 지불한 이후에 최종적으로 당기순이익을 산출한다.

그런데 중소기업은 경영하는 대표가, 혹은 경영하는 타인자본의 부채에 대해서는 당연히 이자를 지급하지만, 자기자본에 대해서는 그 비용이 발생하지 않는 것으로 생각해서는 곤란이 의외로 많다. 하지만 그렇지가 않다. 기업을 경영하는 과정에서 자기자본에 대해서도 비용이 발생하고 그에 대한 배분, 개수, 통틀 수 있어야 안정적 경영을 할 수 있다.

당장을 반박해 자본비용을 출자하는 주주, 즉 투자자 입장에서 생각해도, 출자한 금액에 대하여 타인자본처럼 경영실적과 무관하게 이자를 받을 수 있고, 경영실적이 좋지 않은 경우엔 배당도 받지 못하게 되면, 나아가 회사 전체를 망할 경우 투자 비용 회수에 대한 보장이 되지 않으므로 자기자본에 투자하는 위험률은 타인자본보다 높다고 할 수 있다.

위험률이 높다고 하는 것은 위험이 높은 만큼 요구하는 수익률이 높게 마련이고, 따라서 회사의 자본으로 투자하는 투자자는 타인자본보다 더 높은 수익률을 요구하는 것이 자국의 당연한데, 그에 따라 자기자본 비용은 타인자본 비용보다 더 높은 것이 정상이다.

자본비용은 타인자본 비용과 자기자본 비용을 가중평균하는 방법으로 산출하게 된다. 가중평균자본비용(WACC)을 간단히 설명하면, 현재자본 중 타인자본과 자기자본 비중이 각각 4/10과 6/10이고, 타인자본 비용과 자기자본 비용이 각각 5%, 10%라고 가정할 때,

$$\text{가중평균 자본비용 WACC} = 0.4 \times 5\% + 0.6 \times 10\% = 8\%$$

WACC를 재무담당자가 산출하면 되니 경영자는 몰라도 된다. 다만 WACC가 가중평균 자본비용이라는 뜻만 이해하면 된다. 나아가 직관적으로 더 중요한 것이 차입금에 대한 '책표이자보상비율'이다.

누구나 알고 있는 것처럼 기업은 경영하려면 여러 가지의 자원이 필요하다. 크게 사물로 대표되는 인적·자본·물적 자원으로 나눌 수 있고, 물적 자원의 가장 대표적인 주자가 바로 자본이다.

자본은 무언가 흔히 부채라고 부르는 타인자본과 자기자본으로 나눌 수 있다. 그런데 기업을 경영하는 입장에서 타인자본은 이자를 지급해야만 하고 또한 배분 상환의 의무를 진다. 반면에 자기자본은 이익이 있으면 배당을 하고 그렇지 않으면 배당을 하지 않아도 된다. 나아가 회사 전체가 망할 때까지 상환 의무가 없다. 이런 점에서 자본을 총합하고도 남는 것이 재산이 있다면 그에 대한 배분 할 권리를 가질 뿐이다.

따라서 기업의 경영 상황을 나타내는 재무정보의 표기방식에 있어서도 이런 분별이 적용된다. 기업의 재무상태를 나타내는 '재무상태표'에서도 부채와 자본에 표시되는 무속 회계 표시 순서에 있어서, 타인자본에 대해서만 있다는 의미로 자본보다 부채가 뒤쪽에 표시된다.

일정 기업의 경영실적을 나타내는 '손익계산서'의 경우에도 마찬가지다.

한국경제

한경경영지원단
<http://clean.hankyung.com>

◎ 정보화전략계획수립(ISP) 컨설팅 수행

- 20.05~20.08. 창업진흥원
- 20.10~20.12. 한국연구재단
- 21.01~21.04. 소상공인시장진흥공단
- 21.11~22.06. 서울특별시
- 22.08~22.12. 경찰대학교
- 23.08~24.05 ODA (요르단 경찰청 PSD)

◎ 2022년 AI학습용데이터 구축사업 평가

- | | |
|----|----------------------------|
| 1차 | 08 방송 콘텐츠 대화체 음성인식 데이터 |
| | 09 방송 콘텐츠 한국어·영어 통번역 데이터 |
| | 43 갑각류 종자생산 데이터 |
| | 48 식생 탄소 포집량 식별 데이터 |
| 2차 | 74 축산 기자재(소, 돼지) 3D 데이터 |
| | 75 소(한우, 젃소) 및 돼지 발정행동 데이터 |
| 3차 | 06 인공지능 신기술 선도(자유 공모) |

◎ 데이터 가치평가 컨설팅

- 23.09~23.11 중소벤처기업진흥공단



데이터 관련 강의

- 데이터분석 및 실전 R코딩 (경영기술지도사회, 빅데이터 분석기사 자격증 취득 과정)
- 데이터분석 Python 심화과정 (서울 여성능력개발원 강동 여성인력개발센터 / 용산 여성인력개발센터)
- 파이썬 코딩을 통한 크롤링 자동화 인텐시브 과정 (경영지도사 및 컨설턴트)
- AI & ChatGPT 활용 및 데이터분석 컨설팅 방법론 (경영기술지도사회, 국제공인컨설턴트 CMC 양성과정)
- AI & 데이터분석 (매경아카데미, 동북아 ICT 포럼)





책 쓰기 프로젝트

KYOBObEBook eBook <본의 심리학> 모건 하우절 신작!

2018년 1월

KYOBObEBook 통합검색 > 대입을 결정하는 초등 영어 공부법

대학교재 > 어린이 > 베스트 > 신상품 > 이벤트 > 사운품 > PICKS > CASTing > 교보ONLY

국내도서 > 경제/경영 > 경영전략 > 경영전략일반 >

성장하는 기업의 5가지 조건

한치호, 홍용기, 하현식, 최종철 저(공)

한국경제신문 > 2018년 01월 29일

10.0 (1개의 리뷰) 도움돼요 (100%의 구매자)

“어떻게 지속적으로 성장할 것인가?”

같이 보러봐! 없는 세상엔 시대의 위기를 극복하고
무한 경쟁 시대에 살아남는 전략을 찾아라!

한국경제신문!

챗GPT AI로 데이터분석 마스터하기

초보자도 가능한 노코딩 AI 데이터분석

eBook 6,300원

SAIT 구독

초보자도 가능한
노코딩 AI 데이터분석

챗GPT AI로
데이터분석 마스터하기

홍용기
이원구
홍성일

다즈베즈북스
데이터베이스교육센터

2023년 5월

이 모든 것을 담은
All-In-One
생성형 AI 활용 전략 전자책

ChatGPT 활용 전략

ChatGPT 활용 전략

ChatGPT 활용 전략

2024년 4월

AI와 데이터분석은 전문가들의 전유물이 아니다.
AI & Data Literacy for Everyone.

일반인을 위한
인공지능과 데이터 리터러시

챗GPT-4o(omni) 활용 데이터분석

컨설팅학박사 홍용기 저

일반인들
AI와 데이터의
세계로 안내하는
입문서

인공지능과
데이터분석의
기본개념부터
응용까지

마케팅, 인사,
생산관리, 재무 등
실제 업무 가능할
7가지 분석사례

퍼플

2024년 6월



바로출판
POD


< POD 베스트

- 분야종합
- 소설
- 시/에세이
- 인문
- 가정/육아
- 요리
- 건강
- 취미/실용/스포츠
- 경제/경영
- 자기계발
- 정치/사회
- 역사/문화
- 종교
- 예술/대중문화
- 중/고등참고서
- 기술/공학
- 외국어

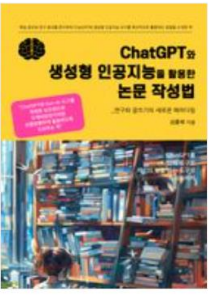
POD 베스트

2024년 06월 ⓘ


월간 ▾




1
일반인을 위한 인공지능과 데이터 리터러시
홍용기 · 퍼플
15,000원



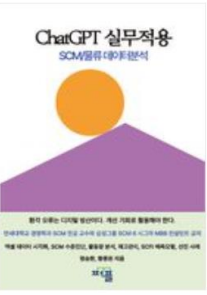
2
ChatGPT와 생성형 인공지능 활용한 논문 작성법
신준석 · 퍼플
46,600원



3
개인정보 보호법 이해와 해설
강문석 · 퍼플
30,000원



4
음원 제작, 이견 알고 하자. 오디오 믹싱 & 마스터링
이대은 · 퍼플
26,900원



5
ChatGPT 실무적용 SCM/물류 데이터분석
정승환, 황종원 · 퍼플
35,000원

초등 어린이 필독서!

이벤트

드래곤볼 GT 세트 30% 할인

쿠폰/혜택

흑자는
인생 뭐 별거 있냐고 할테지만...

인생은 어쩌면
수많은 연결고리들의 집합일지도...

현재를 살면서 긴장을 늦출 수 없는 이유는
이 연결고리가 앞으로 어떻게 연결될지 지금은 알 수 없기 때문이다.

대체 불가능한 사람으로 살아남기 위해 해야 할 것

What to do?



데이터분석 전문가 ADP (Advanced Data Professional)	필기 실기	국가 공인자격	K-Data
빅데이터 분석기사 BAE (Bigdata Analysis Engineer)	필기 실기	국가 기술자격	과기정통부 & 통계청
데이터분석 준전문가 ADsP (Advanced Data Semi Professional)	필기	국가 공인자격	K-Data

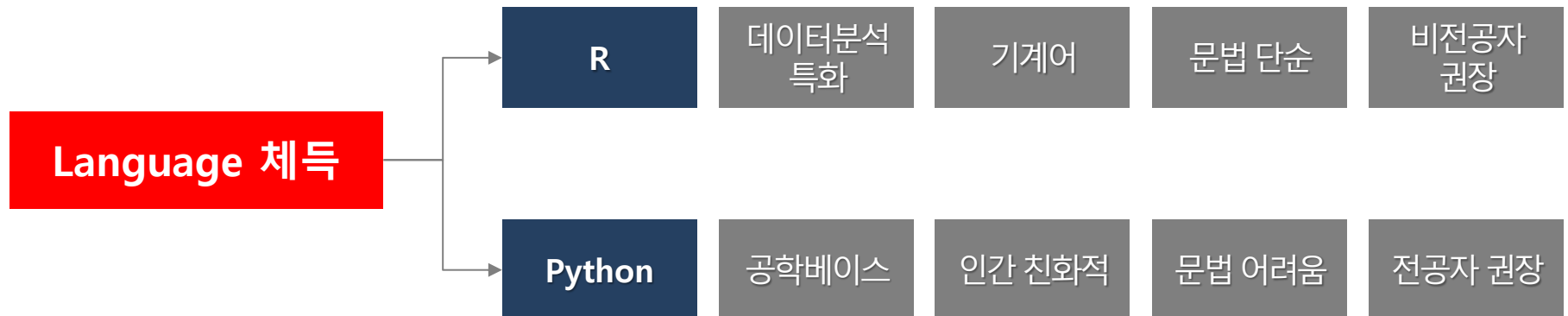
혹자는
인생 뭐 별거 있냐고 할테지만...

인생은 어쩌면
수많은 연결고리들의 집합일지도...

현재를 살면서 긴장을 늦출 수 없는 이유는
이 연결고리가 앞으로 어떻게 연결될지 지금은 알 수 없기 때문이다.

대체 불가능한 사람으로 살아남기 위해 해야 할 것

How to do?





기술의 가치

“가치를 아는 사람에게 기술이 가야 빛을 발한다.”

WWW



가족끼리 암호를 정해 두셨나요?

뉴스속속보

KBS
NEWS

진짜같은 가짜,
딥페이크 직접 제작해보니



Like every great presentation, I've divided my talk into three subjects. Steve Jobs -

I .

**Data &
Information**

II .

**Database &
Schema**

III .

**Data
Gathering**



DATA

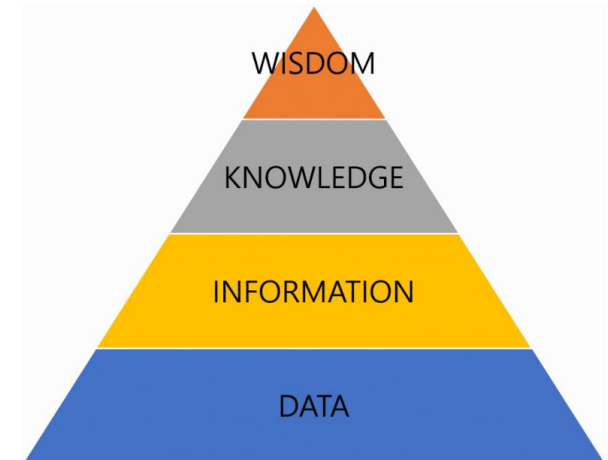
‘추론과 추정의 근거를 이루는 사실’

(A thing given or granted; something known or assumed as fact and made the basis of reasoning or calculation; an assumption of premiss from which inferences are drawn. OED, Vol. IV 264)



Data vs Information

피라미드 요소	설명
지혜 (wisdom)	<ul style="list-style-type: none"> ● 근본 원리에 대한 깊은 이해를 바탕으로 도출되는 창의적 아이디어 ● 상황이나 맥락에 맞게 규칙을 적용하는 요소
지식 (knowledge)	<ul style="list-style-type: none"> ● 다양한 정보를 구조화 하여 유의미한 정보로 분류하고 일반화 시킨 결과물 ● 정보에 기반해 찾아진 규칙
정보 (information)	<ul style="list-style-type: none"> ● 여러가지 데이터 중에 사용자에게 필요한 데이터 ● 사용자의 필요에 따라 정제되거나 가공된 데이터를 정보라고 부름
데이터 (data)	<ul style="list-style-type: none"> ● 객관적 사실로서 가공하기 전의 순수한 자료 ● 가공되지 않은 데이터는 그 자체로는 의미를 지니기 어려움



✓ 전국의 지역별로 조사한 매 시간별 기온 측정자료 → 데이터

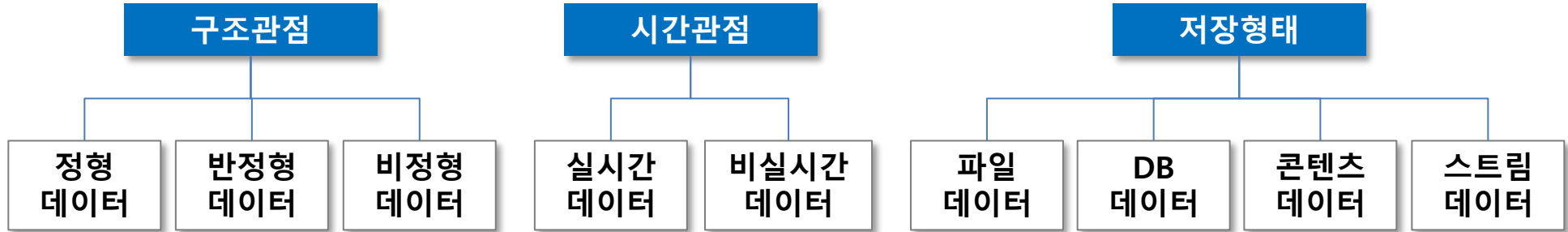
✓ 어떤 목적에 의해 이 데이터를 기반으로 월별 또는 계절별 평균 기온을 산출 → 정보

ex. 어떤 지역의 온도에 따라 혹은 일교차에 따라 어떤 지역에서 어떤 과일을 재배하겠다는 등의 의사결정에 활용

‘데이터분석’ 이라고 하면...

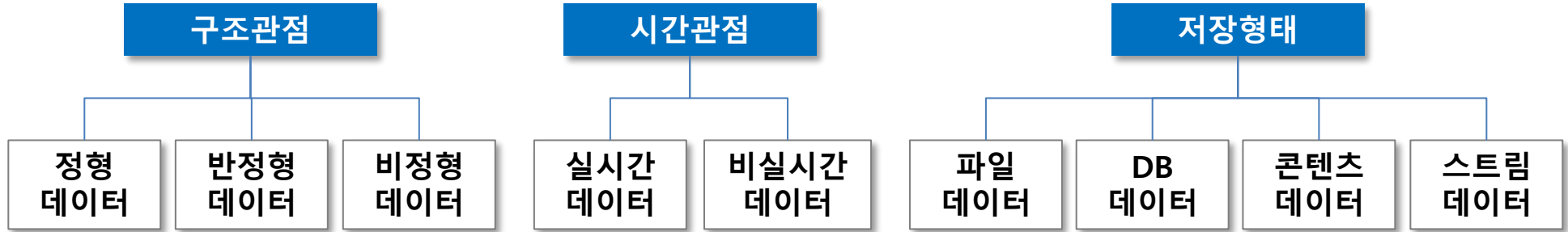
어떠한 데이터를 어느 범위까지
분석해야 하는 것인가?

Type of Data

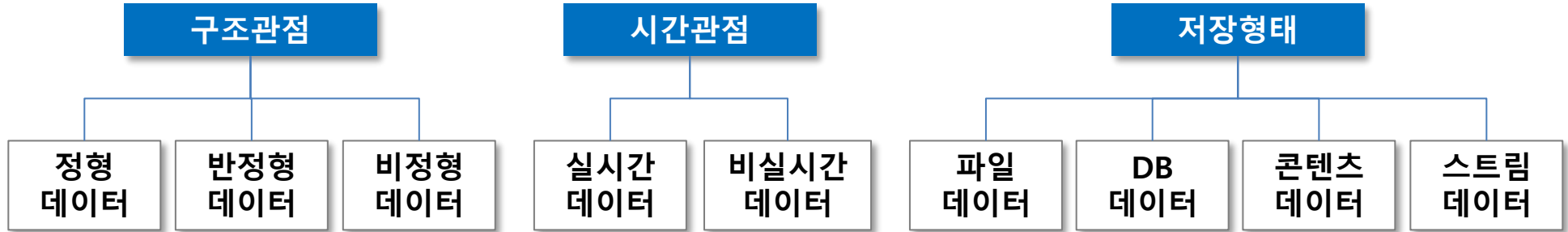


구조관점의 데이터 유형	유형	설명	종류
	정형 데이터	정형화된 스키마(형태) 구조 기반으로 행과 열로 구성	관계형 데이터베이스, 스프레드 시트
	반정형 데이터	스키마 구조 기반이나 값과 형식에서 일관성을 가지지 않는 데이터	XML, HTML, 로그데이터, JSON, 센서데이터
	비정형 데이터	스키마 구조 형태를 가지지 않으며 고정된 필드에 저장되지 않는 데이터	SNS, 게시판, 텍스트, 이미지, 오디오/비디오

Type of Data



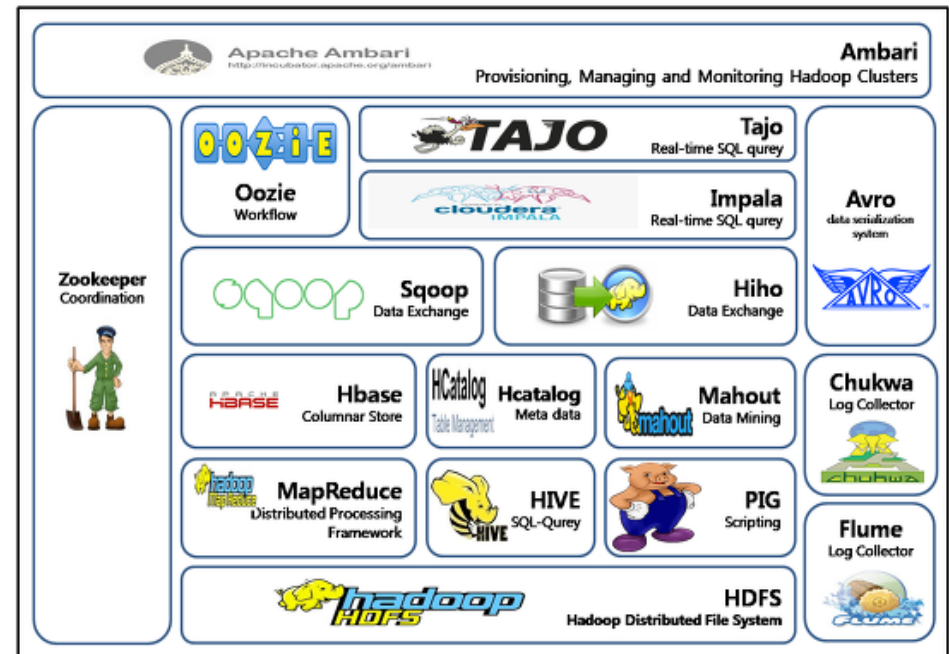
시간관점의 데이터 유형	유형	설명	종류
	실시간 데이터	생성된 이후 수 초 ~ 수 분 이내에 처리되어야 의미가 있는 현재 데이터	센서 데이터, 시스템 로그, 보안 장비 로그, 알람, 네트워크 장비 로그
	비실시간 데이터	이후에 처리되어야 의미가 있는 과거 데이터	통계, 웹 로그, 서비스 로그, 구매정보, 디지털 헬스케어 정보



저장형태 관점의 데이터 유형	유형	설명
	파일 데이터	시스템 로그, 텍스트, 스프레드 시트 등과 같이 파일 형식으로 파일 시스템에 저장되는 데이터
	데이터베이스 데이터	관계형 데이터베이스, NoSQL, 인메모리 데이터베이스 등에 저장된 데이터
	콘텐츠 데이터	텍스트, 이미지, 오디오, 비디오 등과 같이 개별적으로 데이터 객체로 구별 가능한 미디어 데이터
	스트림 데이터	센서 데이터, HTTP 트랜잭션, 알람 등과 같이 네트워크를 통해 실시간으로 전송되는 데이터

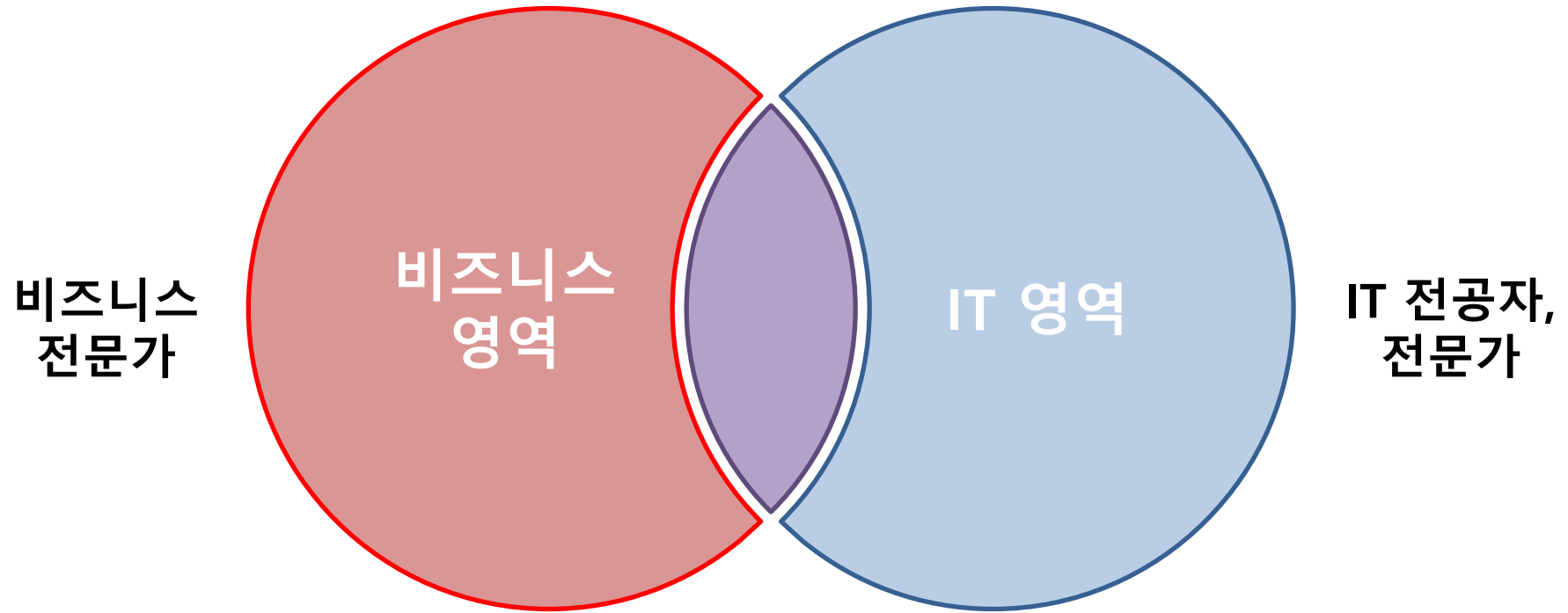
Data Collection Methods and Technologies

<ul style="list-style-type: none"> ETL(Extract, Transform, Load) FTP(File Transfer Protocol) 스쿱(Sqoop) 스크래파이(Scrapy) 	<ul style="list-style-type: none"> 아파치 카프카(Apache Kafka) 플럼(Flume) 스크라이브(Scribe) 척와(Chukwa) 	<ul style="list-style-type: none"> CEP(Complex Event Processing) EAI(Enterprise Application Integration) CDC(Change Data Capture) ODS(Operational Data Store) 	<ul style="list-style-type: none"> 크롤링(Crawling) RSS(Rich Site Summary) Open API 스트리밍(Streaming)
--	---	---	--



※출처: 하둡 에코시스템(Hadoop-Ecosystem)이란 <https://butter-shower.tistory.com/73>





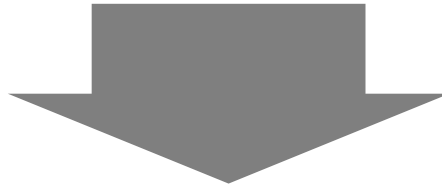
“강사/컨설턴트는 어느 영역까지 cover해야 할 것인가?”



- | | | | | |
|---|--------------------------|----------------------------|-----------------------|-------------|
| ▪ 수십 기가바이트에 이르는
대용량 데이터 기반 | ▪ 데이터 분할 및 전처리 | ▪ 데이터 수치화 및 변수 간의
관계 관찰 | ▪ 분석 모델을 활용한 모델
구축 | ▪ 분석결과 보고 |
| ▪ 데이터 마이닝 수행에
필요한 데이터 선별하여
샘플링 가능 | ▪ 일관성이 없고 불완전한
데이터 정리 | ▪ 시각화 활용 | ▪ 모델의 적합성 평가 | ▪ 시각화 자료 첨부 |
| | ▪ 데이터를 분석하기 좋은
형태로 가공 | | | |

“데이터분석 → 데이터수집 + 실질적인 데이터 분석”

데이터분석 = 데이터수집 + 실질적인 데이터 분석



“어느 것이 더 어려운가?”

어느 것이 더 많은 노력과 시간이 들어가나?



Like every great presentation, I've divided my talk into three subjects. Steve Jobs -

I .

Data &
Information

II .

Database &
Schema

III .

Data
Gathering



일반적으로 엑셀 프로그램을 사용하는 방식

재고관리엑셀프로그램V44.xlsm - Microsoft Excel

파일 홈 삽입 페이지 레이아웃 수식 데이터 검토 보기 개발 도구 추가 기능

(입출고)

새로작성

계산수식

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

일지

출처: https://www.kanghanet/bbs/board.php?bo_table=excel01&wr_id=3750&sca=%EC%9E%AC%EA%B3%A0%EA%B4%80%EB%A6%AC



Primary key, Foreign key

재고관리엑셀프로그램V44.xlsm - Microsoft Excel

파일

홈

삽입

페이지 레이아웃

수식

데이터

검토

보기

개발 도구

추가 기능

(입출고)

새로작성

저장하기

(세트출고)

계산수식

찾기(일지)

(집계현황)

세트출고 작성

세트출고 작성

(재고분석)

집계 수량

집계 금액

(발주서)

기간별수량

기간별금액

(재고분석)

재고분석

모두보기

(재고부족)

재고부족

재고없음

(발주서)

발주서 작성

{발주목록}작성

(발주실적)

발주실적(수량)

발주실적(금액)

(발주서)

발주서

발주서

(분석시트)

월별.분류.품명

입고처.출고처

Kangha.Net

(품목설정)

설정하기

메뉴 명령

C20

가공식품

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
2	입출고 입력														
9		날짜	분류	품명	규격	Code	이전재고량	입고량	출고량	유실량	입고처	출고처	특기사항		
149		2016-07-15	형강	ANGLE	45x45x3tx6m	T016-02	0	5			현대제철				
150		2016-07-15	형강	ANGLE	65x65x6tx10m	T016-03	92	1	55		현대제철	제1공장			
151		2016-07-20	가공식품	간장	샘표	F001-10	52		50			조립반			
152		2016-07-20	스위치	리미트	SD-34	E500-24	142		120			조립반			
153		2016-07-20	철자재	철판(EGI)	3.2*1219*2438	T016-08	39		33			조립반			
154		2016-07-20	형강	환봉	4 22*8m	T016-01	4		2			조립반			
155		2016-07-24	가공식품	고추장	순창	F001-11	0	3			샘표간장				
156		2016-07-24	가공식품	고추장	순창	F001-11	3	5			샘표간장				

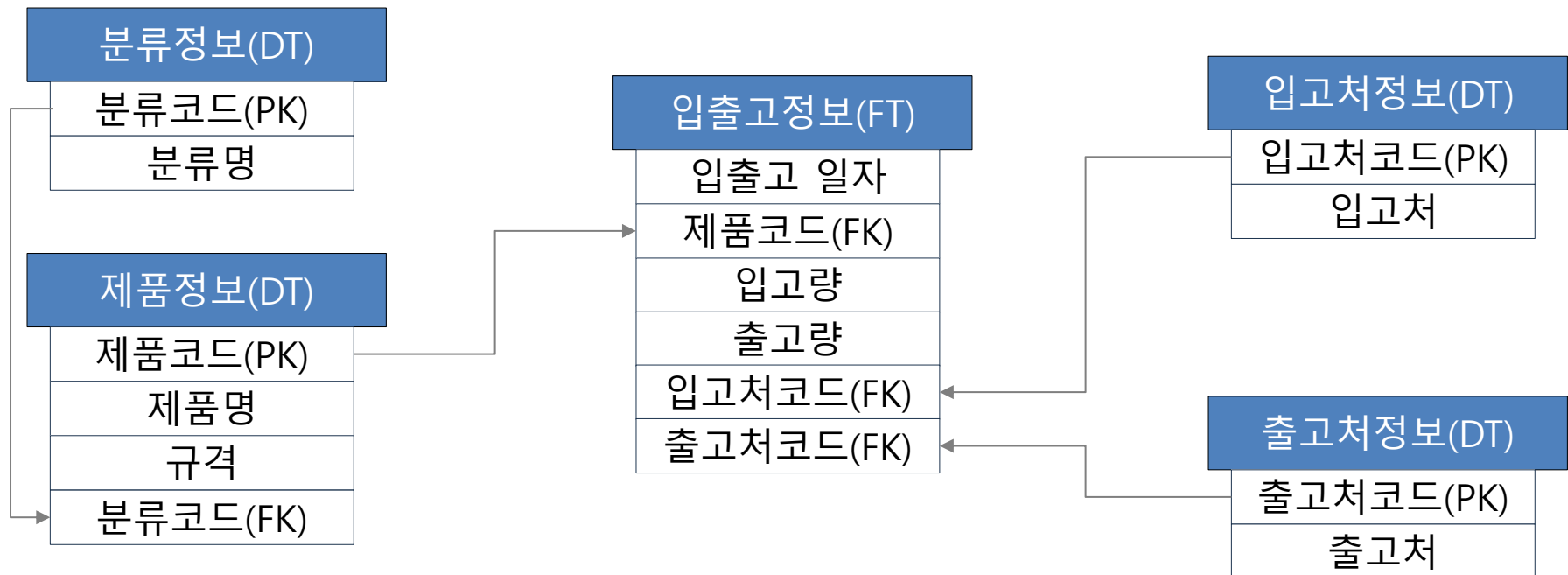
기본 키(Primary key)는 특정 행을 유일하게 식별할 수 있는 속성 (ex. 주민등록번호)

- ✓ Unique한 값으로 중복이 허용되지 않음 → 테이블 생성 시 반드시 하나 이상의 필드/열을 지정
- ✓ Null값 포함이 불가 (Non-null)
- ✓ 가장 세분성(Granulairty)이 높아야 함



데이터 모델링

Dimension 테이블과 Fact 테이블은
특정한 기본키(primary key)와 외래키(foreign key)를 기준으로 상호 연결되며
DT:FT는 일대다(1:n)의 관계(cardinality, 선택도)를 갖게 됨





데이터 모델링

분류코드	분류명
1	형강
2	가공식품
3	스위치
4	철자재
5	센서
6	식품
7	기타

제품코드	제품코드2	제품명	규격	분류코드
1	T016-02	ANGLE	45	1
2	T016-03	ANGLE	65	1
3	F001-10	간장		2
4	E500-24	리미트		3
5	T016-08	철판		4
6	T016-01	환봉		1
7	F001-11	고추장		2
8	E300-04	말굽		5
9	F222-01	김치		6
10	E500-27	푸쉬버튼		3
11	S103-01	파이프	20	4
12	S103-02	파이프	40	4
13	E300-05	온도		5
14	T001-01	페인트		7

입출고일자	제품코드	입고량	출고량	입고처코드	출고처코드
2024-07-15	1	5		1	
2024-07-16	2	1		1	
2024-07-16	2		55		1
2024-07-17	3		50		2
2024-07-18	4		120		2
2024-07-19	5		33		2
2024-07-20	6		2		2
2024-07-21	7	3		2	
2024-07-22	7	5		2	
2024-07-23	7		5		3
2024-07-24	8		3		3
2024-07-25	9	8		3	
2024-07-26	10		5		4
2024-07-27	11	2		4	
2024-07-28	12		11		1
2024-07-29	13	5		4	
2024-07-30	3		1		5
2024-07-31	14		3		5

입고처코드	입고처
1	현대제철
2	샘표간장
3	청정원
4	포스코

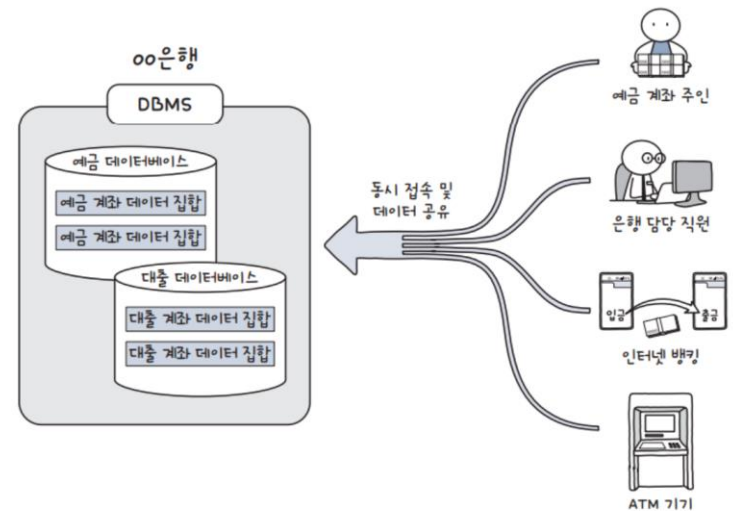
출고처코드	출고처
1	제1공장
2	조립반
3	제2공장
4	금형반
5	제3공장

※ 참고: Star 스키마와 Snowflake 스키마 <https://www.integrate.io/ko/blog/snowflake-schemas-vs-star-schemas-what-are-they-and-how-are-they-different-ko/>



데이터베이스

- **데이터베이스**(database, DB): 여러 사람이 공유하여 사용할 목적으로 체계화해 통합, 관리하는 **데이터의 집합**이다. (위키백과)
- **DBMS**(Database Management System): 데이터베이스를 운영하고 관리하는 소프트웨어(예: MySQL, Oracle, SQL 서버, PostgreSQL, MariaDB)
- **SQL**(Structured Query Language): 구조화된 질의 언어라는 뜻으로 관계형 데이터베이스에서 사용되는 언어
- 데이터베이스는 1개 이상의 **테이블**로 이루어지며, 테이블은 **fact 테이블**과 **dimension 테이블**로 구분됨



※그림 출처: 혼공러들의스터디공간 <https://honggongharibokk/%EB%8D%B0%EC%9D%B4%ED%84%B0%EB%B2%A0%EC%9D%B4%EC%8A%A4%EC%9D%B4%ED%95%B4%ED%95%B8%EA%B8%B0-database-dbms.sql%EC%9D%B8%EA%B0%9C%BB%80>



데이터베이스

```
Microsoft Windows [Version 10.0.22631.3155]
(c) Microsoft Corporation. All rights reserved.

C:\Users\821033669010>cd "C:\program files\MariaDB 10.8\bin

C:\Program Files\MariaDB 10.8\bin>mysql.exe -u root -p
Enter password: *****
Welcome to the MariaDB monitor.  Commands end with ; or \g.
Your MariaDB connection id is 6
Server version: 10.8.5-MariaDB mariadb.org binary distribution

Copyright (c) 2000, 2018, Oracle, MariaDB Corporation Ab and others.

Type 'help;' or '\h' for help. Type '\c' to clear the current input
statement.

MariaDB [(none)]> SHOW databases;
+-----+
| Database |
+-----+
| financial_terms |
| information_schema |
| investar |
| kotra |
| mysql |
| opentutorials |
| performance_schema |
| scraping |
| sys |
| wikipedia |
| workbench |
+-----+
11 rows in set (0.002 sec)
```

```
MariaDB [(none)]> USE opentutorials;
Database changed
MariaDB [opentutorials]> SHOW tables;
+-----+
| Tables_in_opentutorials |
+-----+
| author |
| topic |
| topic_backup |
+-----+
3 rows in set (0.010 sec)
```

```
MariaDB [opentutorials]> DESC topic;
+-----+-----+-----+-----+-----+-----+
| Field | Type | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| id | int(11) | NO | PRI | NULL | auto_increment |
| title | varchar(100) | NO | | NULL | |
| description | text | YES | | NULL | |
| created | datetime | NO | | NULL | |
| author_id | int(11) | NO | | NULL | |
+-----+-----+-----+-----+-----+-----+
5 rows in set (0.018 sec)
```

```
MariaDB [opentutorials]> SELECT * FROM topic;
+-----+-----+-----+-----+-----+
| id | title | description | created | author_id |
+-----+-----+-----+-----+-----+
| 1 | MySQL Database | MySQL is Relational ... | 2023-12-19 10:03:32 | 1 |
| 2 | Oracle | Oracle is ... | 2023-12-19 10:09:53 | 1 |
| 5 | MongoDB | MongoDB is ... | 2023-12-19 11:29:44 | 1 |
+-----+-----+-----+-----+-----+
3 rows in set (0.011 sec)
```



Like every great presentation, I've divided my talk into three subjects. Steve Jobs -

I .

Data &
Information

II .

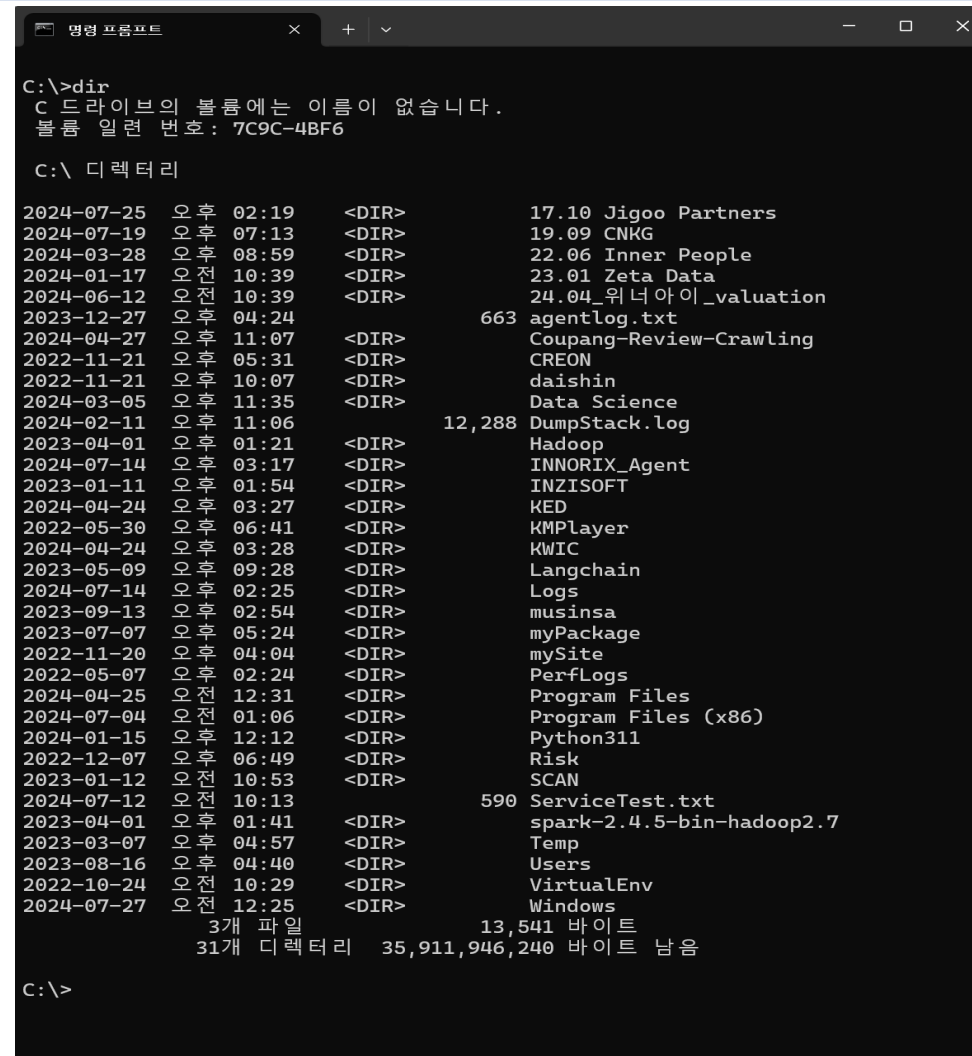
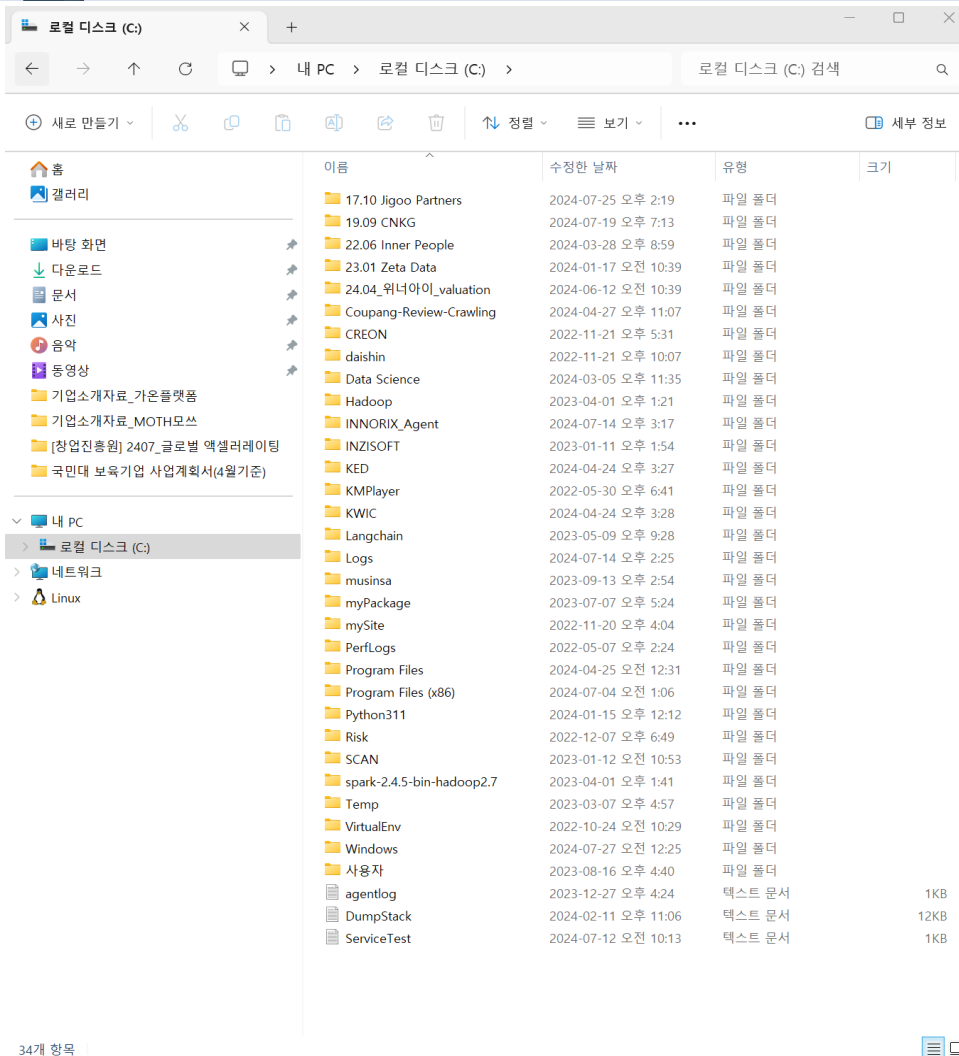
Database &
Schema

III .

**Data
Gathering**



Disk Operation System





웹 데이터 가져오기

단 3줄로 가능한 쿠팡 상품 댓글 크롤링

1. 인용 사이트 : <https://github.com/JaehyoJJAng/Coupang-Review-Crawling>
2. Git 설치 (<https://git-scm.com/downloads/win>)
→ 참조 사이트 : <https://m.blog.naver.com/insideimagix/223325204776>
3. Python 설치
→ (<https://dotiromooook.tistory.com/32>)
4. 명령 프롬프트 또는 “cmd” 실행 → DOS 화면 뜸
5. \$ git clone https://github.com/JaehyoJJAng/Coupang-Review-Crawling.git
6. \$ cd Coupang-Review-Crawling
7. \$ pip install -r ./requirements.txt

반드시
참조

실행1 \$ python main.py

실행2 복사된 url 주소 붙여넣기 → Enter → 페이지 수 지정 → Enter



MS Excel vs Google Sheets



Google Sheets