

# 데이터베이스와 SQL

1장

## 데이터베이스 소개

빅데이터 분석가 과정

# 목차

- 데이터베이스 소개
- SQL
- MySQL

# 데이터 베이스 소개

## ■ 데이터베이스

- ‘데이터의 집합’
- 여러 사용자나 응용 프로그램이 공유하고 동시에 접근 가능한 데이터
- ‘데이터의 저장 공간’ 자체를 의미하기도 함

## ■ DBMS (Database Management System)

- 데이터베이스를 관리·운영하는 소프트웨어
- 사용자나 응용 프로그램은 DBMS가 관리하는 데이터에 동시에 접속하여 데이터를 공유함

# 데이터베이스 소개

## ■ DBMS 개념도

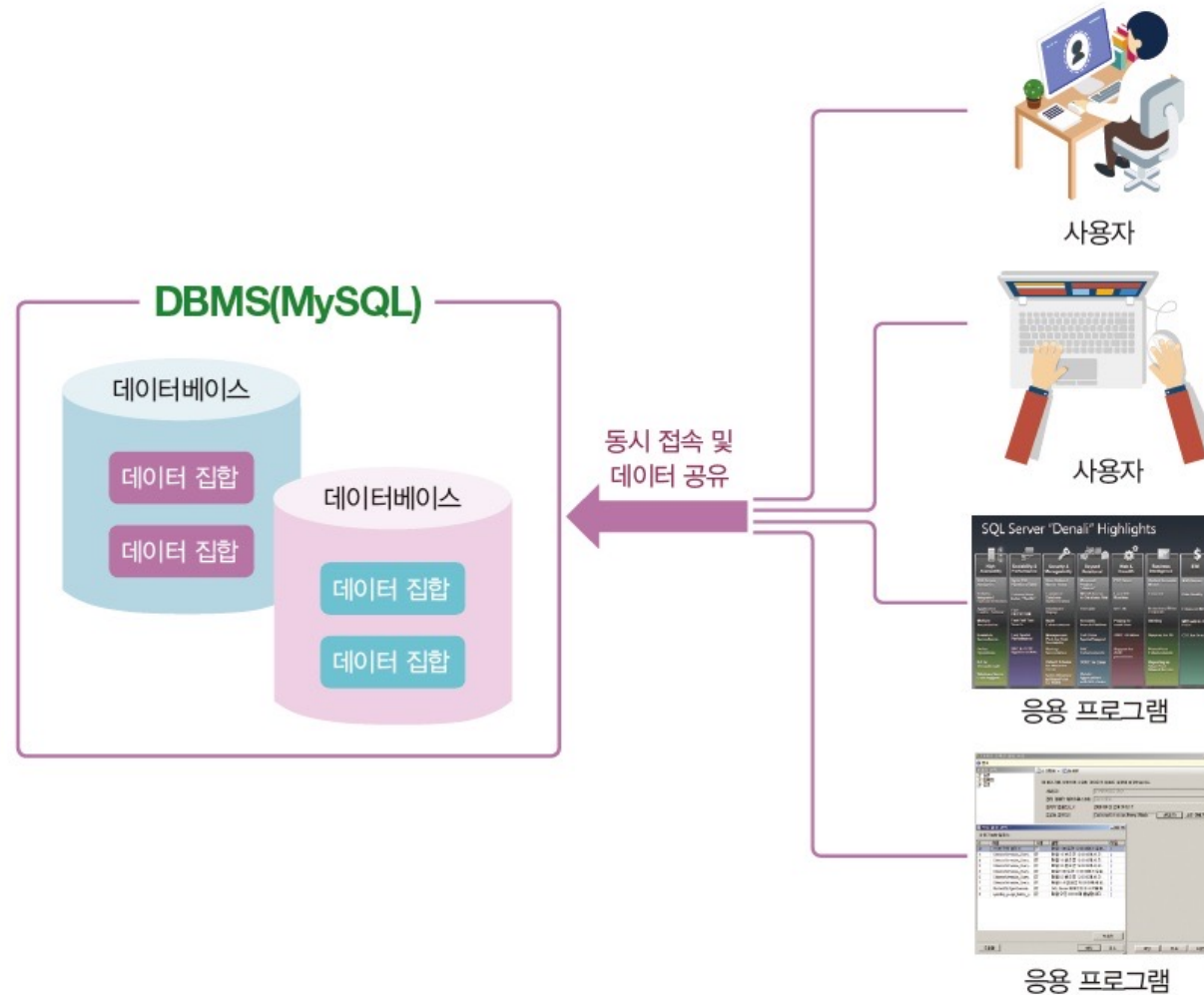


그림 1-1 DBMS의 개념

# 데이터베이스의 정의와 특징

## ■ 데이터베이스의 특징

### ■ 데이터의 무결성

- 데이터베이스 안의 데이터는 어떤 경로를 통해 들어왔든 오류가 없어야 함

### ■ 데이터의 독립성

- 데이터베이스와 응용 프로그램은 서로 의존적인 관계가 아니라 독립적인 관계임

### ■ 보안

- 데이터베이스 안의 데이터는 데이터를 소유한 사람이나 데이터에 접근이 허가된 사람만 접근할 수 있음

### ■ 데이터 중복 최소화

- 데이터베이스에서는 동일한 데이터가 여러 군데 중복 저장되는 것을 방지함

### ■ 응용 프로그램 제작 및 수정 용이

- 통일된 방식으로 응용 프로그램을 작성할 수 있고 유지·보수 또한 쉬움

### ■ 데이터의 안전성 향상

- 데이터가 손상되는 문제가 발생하더라도 원래의 상태로 복원 또는 복구할 수 있음

# 데이터베이스 발전

## ■ 오프라인 관리

- 종이에 기록해 장부로 관리



## ■ 파일 시스템 사용

- 컴퓨터 파일에 기록/저장: 엑셀 등 활용
- 데이터의 양이 많아지면 데이터의 관리가 어려움
- 데이터 중복 발생 가능

userID	name	birthYear	addr	mobile1	mobile2	height	mDate
BBK	바비킴	1973	서울	010	0000000	176	2013-05-05
EJW	은지원	1972	경북	011	8888888	174	2014-03-03
JKW	조관우	1965	경기	018	9999999	172	2010-10-10
JYP	조용필	1950	경기	011	4444444	166	2009-04-04
KBS	김범수	1979	경남	011	2222222	173	2012-04-04
KKH	김경호	1971	전남	019	3333333	177	2007-07-07
LJB	임재범	1963	서울	016	6666666	182	2009-09-09
LSG	이승기	1987	서울	011	1111111	182	2008-08-08
SSK	성시경	1979	서울	NULL	NULL	186	2013-12-12
YJS	윤종신	1969	경남	NULL	NULL	170	2005-05-05

# 데이터베이스 발전

- 데이터베이스 관리시스템
  - 파일시스템의 단점 보완
  - 대용량의 데이터를 보다 효율적으로 관리하고 운영하기 위해 사용
  - 데이터의 집합인 데이터베이스를 잘 관리하고 운영하기 위한 시스템 또는 소프트웨어
- SQL(Structured Query Language)
  - DBMS에 데이터 구축/관리/활용을 위해 사용되는 언어
  - DBMS를 통해 중요한 정보들을 입력, 관리, 추출

# 관계형 모델

## ■ 관계형 DBMS

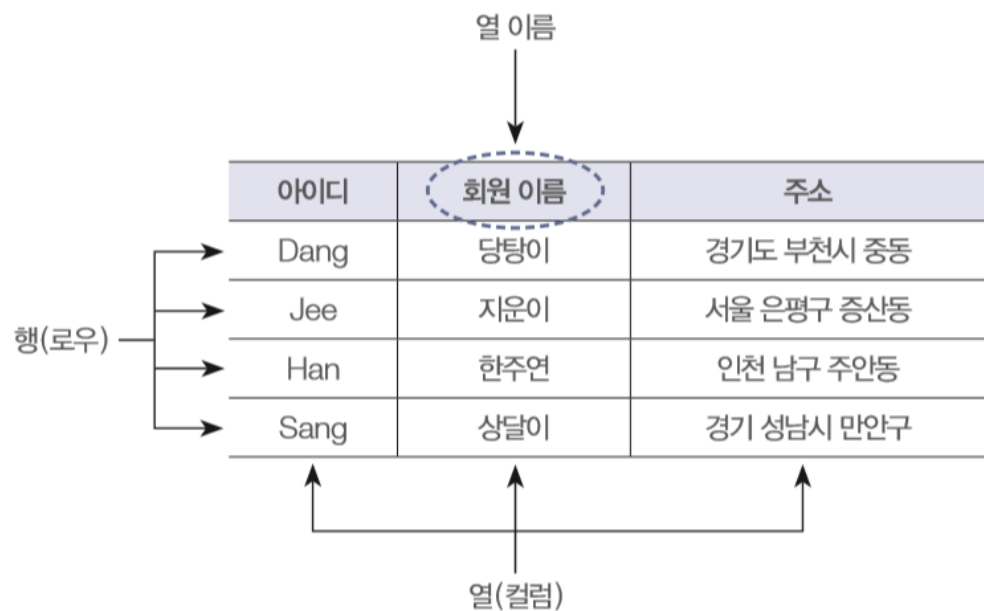
- 모든 데이터는 테이블에 저장
- 테이블 간의 관계는 기본키(PK)와 외래키(FK)를 사용하여 맺음(부모-자식 관계)

## ■ 장점

- 다른 DBMS에 비해 업무 변화에 따라 바로 순응
- 유지·보수 측면에서 편리
- 대용량 데이터를 체계적으로 관리할 수 있음
- 데이터의 무결성(integrity)이 잘 보장됨
  - 정확한 데이터가 유지됨을 보장

## ■ 단점

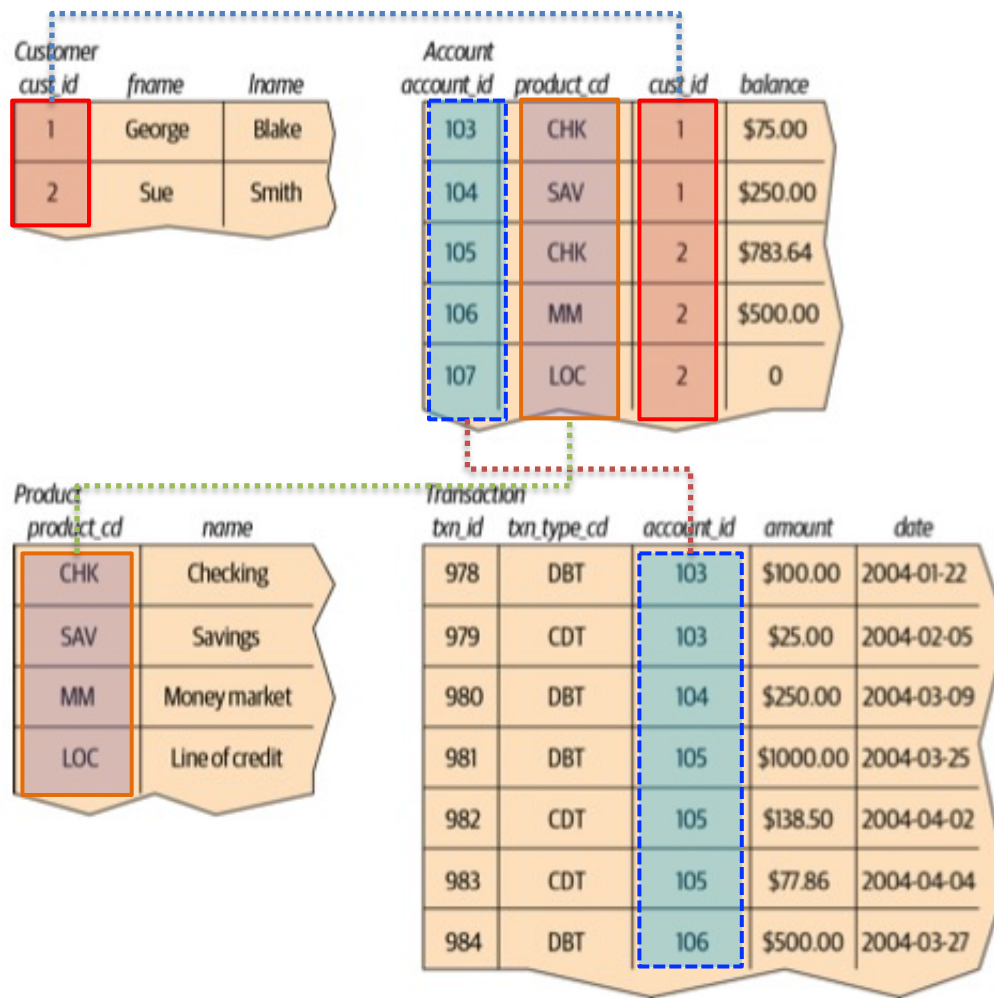
- 시스템 자원을 많이 차지
- 시스템이 전반적으로 느려지는 단점이 있음 (하드웨어 발전으로 해결)





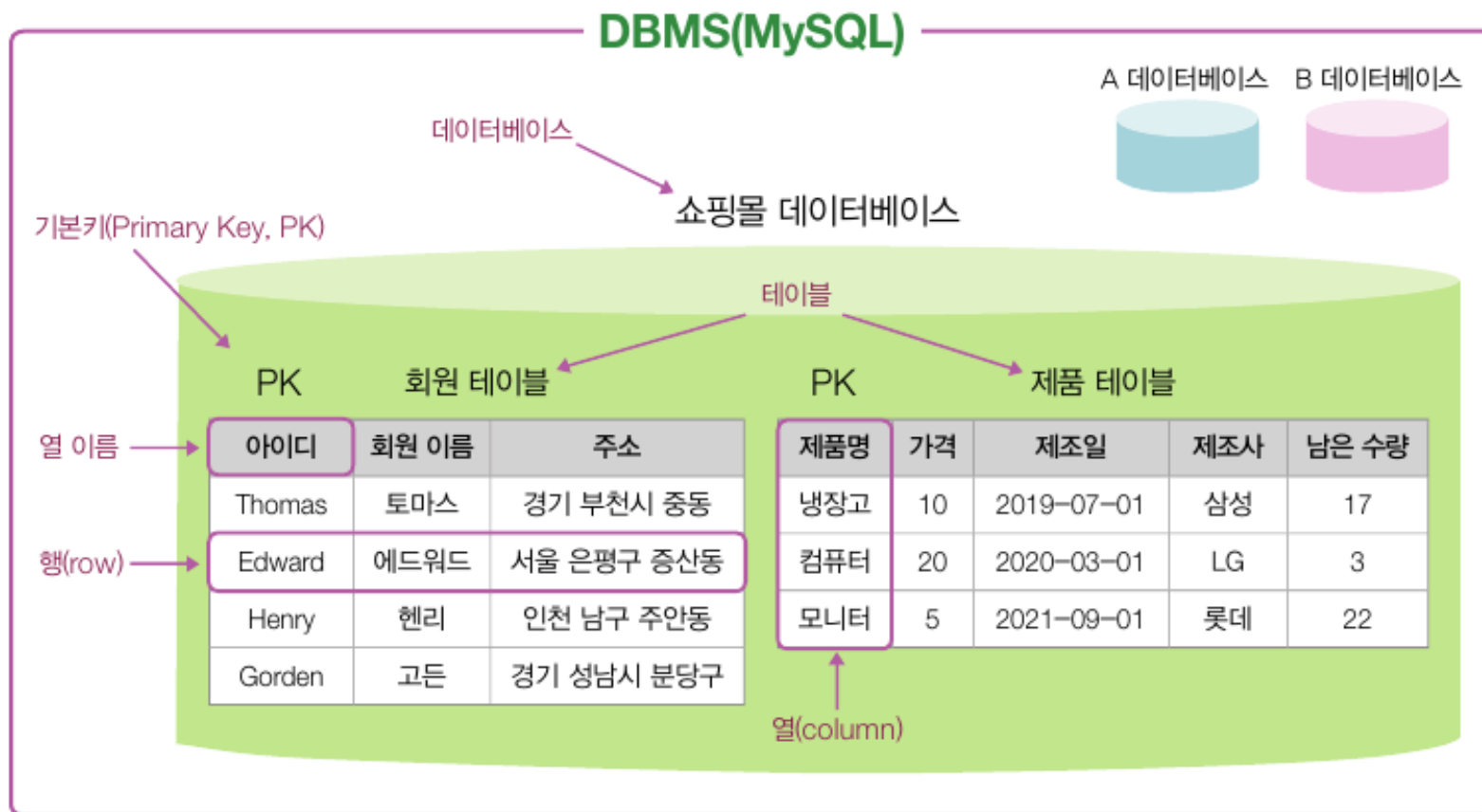
# 관계형 모델

## ■ 계좌 데이터의 관계형 View



# 주요 용어

## ■ 테이블의 구조와 관련 용어



# 주요 용어

## ■ 주요 용어와 정의

용 어	정 의
엔티티(entity)	데이터베이스에서 사용자들이 관심을 갖는 모든 요소 (고객, 부품, 위치정보 등)
열(column, field)	테이블에 저장된 개별적인 데이터 조각 ( <b>필드</b> )
열이름	각 열을 구분하기 위한 이름 각 테이블 내에서는 중복되지 않고 고유해야 됨
데이터 형식	열의 데이터 형식 테이블을 생성할 때 열 이름과 함께 지정
행(row, record)	실질적인 데이터 ( <b>레코드</b> 라고도 부름)
테이블(table)	데이터를 입력하기 위해 표 형태로 만든 것, 행의 집합
결과셋(result set)	SQL 쿼리의 결과물
기본 키(primary key)	각 행을 구분하는 고유 식별자, <b>기본 키는 중복되거나 비어 있으면 안됨</b>
외래 키(foreign key)	다른 테이블에서 행을 구분하는 식별자, <b>두 테이블을 연결하는 key</b>
정규화(normalization)	외래키를 제외하고 각각의 독립적인 정보가 한 위치에만 저장되도록 데이터베이스를 설계하고 수정하는 절차
조인(join)	2개 이상의 테이블을 묶어서 하나의 결과 테이블을 만드는 것

# SQL의 개요

## ■ SQL(Structured Query Language, Sequel)

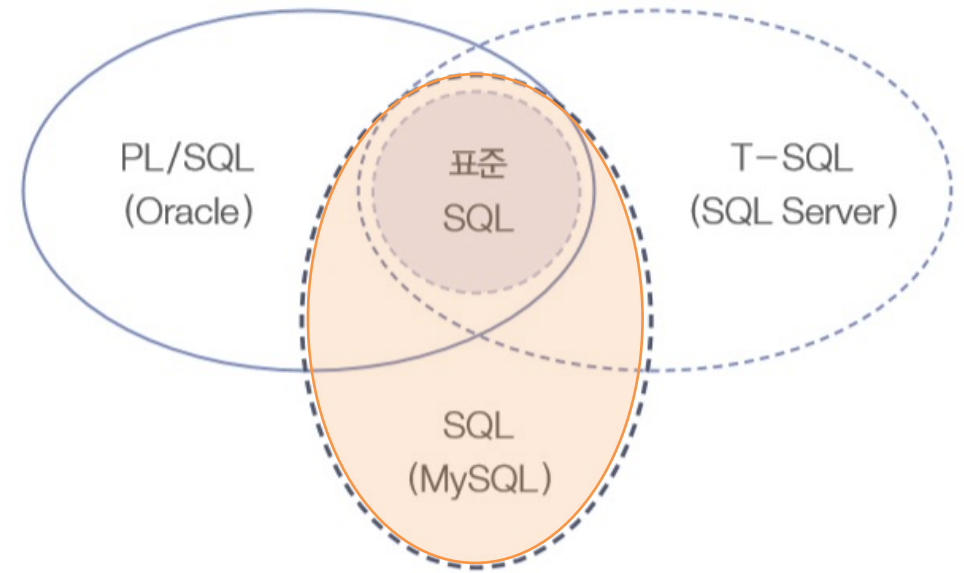
- 데이터베이스를 조작하는 언어
- 구조화된 질의 언어

## ■ SQL의 특징

- DBMS 제작 회사와 독립적임
- 다른 시스템으로의 이식성이 좋음
- 표준이 계속 발전함
- 대화식 언어임
- 클라이언트/서버 구조 지원함

## ■ 표준 SQL과 각 회사의 SQL

- 많은 회사가 되도록 표준 SQL을 준수하려고 노력
- 각 회사의 DBMS마다 특징이 있기때문에 현실적으로 완전히 통일되기는 어려움
- 각 회사의 제품은 모두 표준 SQL을 공통으로 사용하면서 자기 제품의 특성에 맞춘 호환되지 않는 SQL 문 사용



# SQL 스키마

## ■ SQL 스키마(SQL Schema)

- 데이터베이스의 구조와 제약조건에 관해 **전반적인 명세를 기술한 것**
- 속성, 개체, 관계에 대한 정의

```
# 한 줄 주식
/* 여러 줄 주식
 * corporation 테이블 생성
 * corp_id, name 컬럼 추가
 */
create table corporation
(corp_id smallint,
 name VARCHAR(30),
 constraint pk_corporation primary key(corp_id)
);
```



**Questions?**