

# 데이터 입출력 구현

이 자료는 주말코딩(youtube.com/@weekendcode)의 자료입니다.

무단 반출 및 복제 배포를 금지합니다.

## 데이터 모델링 순서

데이터베이스 계획 → 요구사항 분석 → 개념적 데이터 설계 → 논리적 데이터 설계 → 물리적 데이터 설계

### 개념적 데이터 모델링?

- 개체 타입, 속성 등을 명시해서 현실세계를 반영

### 논리 데이터 모델링?

- 개념적 구조를 정규화하고, 규칙과 관계를 완성
- (엔티티, 속성, 관계를 구조적으로 정의)

### 물리 데이터 모델링?

- 레코드 양식 순서, 경로 인덱싱, 클러스터링, 해싱 등

# 관계형 데이터베이스

## 관계형 데이터베이스란?

- 데이터베이스라고 불리는 개념으로 엑셀과 비슷한 "테이블" 형태로 데이터 제공

## 테이블 = 릴레이션(relation)

- 속성(attribute)
- 도메인(domain)
- 튜플(tuple)

예시

학번	이름	핸드폰	이메일
1234	홍길동	010-1234-3287	hong@email.com
5678	박문수	011-3333-2222	park@email.com
9012	이순신	010-9999-2222	lee@email.com

- 차수(Degree): 속성 수 = 4
- 카디널리티(Cardinality): 튜플 수 = 3

- 내포(Intension), 외포(Extension)

---

## 도메인이란?

---

- 속성이 값을 가질 수 있는 범위  
예: Color = 문자열, Year = 정수형

예시

학번	이름	나이	성별	강의명
101	홍길동	23	남	프로그래밍1
102	나미리	21	여	자료구조
103	김미양	25	여	웹 보안

---

## ⚠ 이상(Anomaly)

---

데이터베이스에서 조작할 때 비정상적으로 동작하는 현상

### 삽입 이상

- 새로운 데이터를 삽입할 때 의도하지 않은 데이터까지 삽입됨
- 예시: 자료를 삽입하려는데, 의도하지 않은 자료까지 삽입해야 정상적 반영)

### 갱신 이상

- 중복된 데이터 중 일부만 수정되어 모순이 발생
- 예시: 특정 데이터를 수정했는데, 데이터가 중복 상태에서 하나만 변경됨

### 삭제 이상

- 의도한 데이터 외 다른 정보도 함께 삭제됨
- 예시: 특정 속성을 삭제했는데, 해당 속성을 가진 데이터도 함께 삭제됨

---

## 정규화

---

테이블 구조를 체계화하여 중복을 없애고 무결성을 확보

(주의! 반정규화(역정규화)라는 개념도 있음: 중복을 일으켜서 성능을 향상시킴)

### ◇ 제1 정규형 (1NF)

- 모든 속성이 원자값(더 이상 나눌 수 없는 값)

## ◇ 제2 정규형 (2NF)

- 1NF 만족 + 부분 함수적 종속 제거

## ◇ 제3 정규형 (3NF)

- 2NF 만족 + 이행적 종속 제거

## ◇ 보이스-코드 정규형 (BCNF)

- 3NF 만족 + 모든 결정자가 후보키

## ◇ 제4 정규형 (4NF)

- BCNF 만족 + 다치 종속 제거

## ◇ 제5 정규형 (5NF)

- 4NF 만족 + 조인 종속 제거 (테이블을 projection으로 재구성 가능)

## 종속에 대하여

### ◇ 다치 종속 (Multivalued Dependency)

- 하나의 속성에 여러 개가 매핑될 때

학생이름	수강과목	동아리
홍길동	수학	사진
홍길동	컴퓨터공학	건담조립

$X \twoheadrightarrow Y$

### ◇ 함수적 종속 (Functional Dependency)

- 하나를 알면, 다른 것도 알 수 있다

고객번호	주민등록번호
13345	950212-1000000
13346	981203-2000000

$X \rightarrow Y$

### ◇ 조인 종속 (Join Dependency)

- 여러 테이블을 조합했을 때 결과로 구성 가능

$JD(R_1, R_2, \dots, R_n)$

## 그 외 용어 정리

## 1. 트랜잭션 특성과 4가지

- 원자성, 일관성, 독립성(격리성), 지속성(영속성)

## 2. 샤딩

- 대규모 DB를 여러 독립적인 부분으로 분할해 성능 향상

## 3. 인덱스

- 검색 속도 향상을 위한 추가 저장 구조

## 4. 시스템 카탈로그

- DB에 저장된 데이터 개체에 대한 정보를 저장한 시스템

## 5. 분산 데이터베이스 목표

- 위치 투명성, 중복 투명성, 병행 투명성, 장애 투명성

## 6. 데이터베이스 회복기법

- 즉시갱신, 지연갱신, 검사시점, 그림자 페이징, 미디어 회복기법