

데이터 정규화

데이터 정규화

잘못 설계된 스키마를
더 작은 속성들로 쪼개어
바람직한 스키마가 될 수 있도록 만들어 가는 과정

종속성

잘못 설계된 스키마를 분류하는 기준점

Anomaly

이상현상

삽입이상

삭제이상

갱신이상

비정규 Relation

제 1 정규형

도메인이 원자값

제 2 정규형

부분 함수적 종속 제거

제 3 정규형

이행 함수적 종속 제거

BCNF

결정자이면서 후보키가 아닌 것 제거

제 4 정규형

다치 종속 제거

제 5 정규형

조인 종속성 이용

비정규 Relation

제 1 정규형

도메인이 원자값

제 2 정규형

부분 함수적 종속 제거

제 3 정규형

이행 함수적 종속 제거

BCNF

결정자이면서 후보키가 아닌 것 제거

제 4 정규형

다치 종속 제거

제 5 정규형

조인 종속성 이용

각 정규형은 이전 단계 정규형을 무조건 만족해야 함

비정규 Relation

제 1 정규형

도메인이 원자값

제 2 정규형

부분 함수적 종속 제거

제 3 정규형

이행 함수적 종속 제거

BCNF

결정자이면서 후보키가 아닌 것 제거

제 4 정규형

다치 종속 제거

제 5 정규형

조인 종속성 이용

01

제 1 정규화

회원번호	회원이름	프로그램
101	민지	스쿼시
102	해린	헬스
103	하니	헬스, 테니스

01

제 1 정규화

회원번호	회원이름	프로그램
101	민지	스쿼시
102	해린	헬스
103	하니	헬스
104	하니	테니스

회원번호	회원이름	프로그램
101	민지	스쿼시
102	해린	헬스
103	하니	헬스
104	하니	테니스

제 1 정규화
한 칸엔 하나의 데이터만

02 제 2정규화

회원번호	회원이름	프로그램	가격	납부여부
101	민지	스쿼시	5000	0
102	해린	헬스	6000	0
103	하니	헬스	6000	1
104	하니	테니스	8000	1

02 제 2정규화

회원번호	회원이름	프로그램	가격	납부여부
101	민지	스쿼시	5000	0
102	해린	헬스	8000	0
103	하니	헬스	8000	1
104	하니	테니스	8000	1

02 제 2정규화

회원번호	회원	프로그램	가격	납부여부
101	민	스쿼시	5000	0
102	해	헬스	8000	0
103	하	테니스	8000	1
104	하니	테니스	8000	1

회원번호

회원이름

프로그램

가격

납부여부

101

민지

스쿼시

5000

0

102

민지

스쿼시

5000

0

103

하니

테니스

8000

1

104

하니

테니스

8000

1

제 2 정규화

현재 테이블의 주제와 관련없는 컬럼을
다른 테이블로 빼내는 작업

수업명	강사	출신그룹	가격
발성의 이해	정국	BTS	2000
KPOP 개론	사쿠라	르세라핌	800
AI 커버 만들기	하니	뉴진스	1500
무용학개론	하니	뉴진스	500

중속



수업명	강사	출신그룹	가격
발성의 이해	정국	BTS	2000
KPOP 개론	사쿠라	르세라핌	800
AI 커버 만들기	하니	뉴진스	1500
무용학개론	하니	뉴진스	500

수업명	강사	가격
발성의 이해	정국	2000
KPOP 개론	사쿠라	800
AI 커버 만들기	하니	1500
무용학개론	하니	500

강사	출신그룹
정국	BTS
사쿠라	르세라핌
하니	뉴진스

수업명	강사	가격
발성의 이해	정국	2000
KPOP 개론	사쿠라	1000
AI 커버 만들기	하니	1500
무용학개론	하니	500

강사	출신그룹
정국	BTS
사쿠라	르세라핌
하니	뉴진스

제 3 정규화

일반 컬럼에만 종속된 컬럼은
다른 테이블로 빼내기

데이터 정규화의 장점

각종 이상 현상들이 발생하는 문제점을 해결

일관성, 정확성, 단순성, 비중복성, 안정성 UP

동일한 의미의 일반 속성이 하나의 테이블로 집약됨

=> 한 테이블의 데이터 용량이 최소화되는 효과

데이터 정규화의 단점

릴레이션의 분해로 인해 릴레이션 간의 연산(JOIN 연산)이 많아짐

특히 **조회**를 하는 SQL 문장에서 조인이 많이 발생하여
이로 인한 성능저하가 나타나는 경우에 **반정규화가 필요**