

속전속결
한번에 합격하는
SQLD

4

성능 데이터모델링 개요, 정규화 및 반정규화와 성능

1 학습목표

- ◆ 성능 데이터 모델링 정의와 고려사항에 대해 알아본다.
- ◆ 정규화에 대한 개념과 방법에 대해 알아본다.
- ◆ 반정규화에 대한 개념과 방법에 대해 알아본다.



1 성능 데이터 모델링 정의

DB 성능향상(조회 혹은 데이터 입력/수정/삭제)을 목적으로
설계단계의 데이터 모델링 때부터
정규화, 반정규화, 테이블통합, 테이블분할, 조인구조, PK, FK 등
여러 성능과 관련된 사항이 데이터모델링에 반영될 수 있도록 하는 것



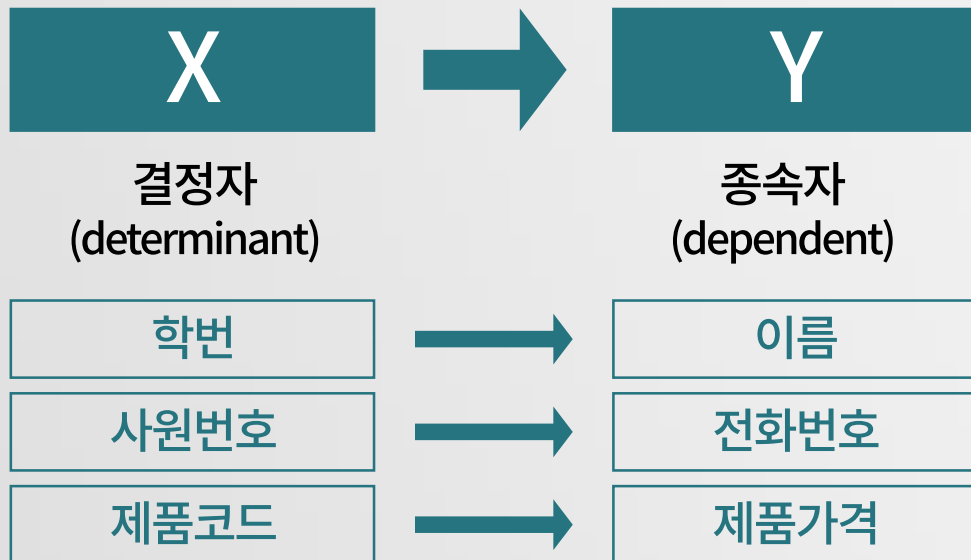
2 고려사항 순서

- ◆ 데이터 모델링 정규화 정확하게 수행
- ◆ DB 용량산정 수행
- ◆ DB 발생하는 트랜잭션 유형 파악
- ◆ 용량,트랜잭션 유형에 따라 반정규화 수행
- ◆ 이력모델 조정, PK/FK 조정, 슈퍼/서브타입 조정
- ◆ 성능관점에서 데이터 모델 검증



1 함수적 종속성

데이터들이 어떤 기준값에 의해 종속되는 현상.
정규화 수행 시, 함수적 종속성에 근거하여 수행하게 됨.



2 정규화 개념

반복적(중복적)인 데이터를 분리하고,
각 데이터가 종속된 테이블에 적절하게 배치되도록 하는 것

거래 번호	제품 코드	제품명	수량	가격
A001	01	아이스 아메리카노	1	4,000
A001	02	그린티 라떼	1	6,000
A002	01	아이스 아메리카노	1	4,000
A003	01	아이스 아메리카노	2	8,000
A004	02	그린티 라떼	1	6,000

정
규
화

거래번호	제품코드	수량	가격
A001	01	1	5,000
A001	02	1	6,000
A002	01	1	5,000
A003	01	2	10,000
A004	02	1	6,000

제품코드	제품명
01	아이스 아메리카노
02	그린티 라떼

2 정규화 개념

반복적(중복적)인 데이터를 분리하고,
각 데이터가 종속된 테이블에 적절하게 배치되도록 하는 것

정규화된 데이터모델	
입력	성능 향상
수정	성능 향상
삭제	성능 향상
조회	성능 향상 혹은 저하



3 1차 정규화

모든 속성은 반드시 하나의 값을 가져야 한다.
(속성값이 원자성을 확보해야 한다.)

제1정규형 위배

제품코드	색상	가격
A01	녹색, 파랑	1,000
A02	빨강	5,000
A03	빨강, 노랑	7,000
A04	파랑	3,000

제품코드	가격
A01	1,000
A02	5,000
A03	7,000
A04	3,000

제품코드	색상
A01	녹색
A01	파랑
A02	빨강
A03	빨강
A03	노랑
A04	파랑

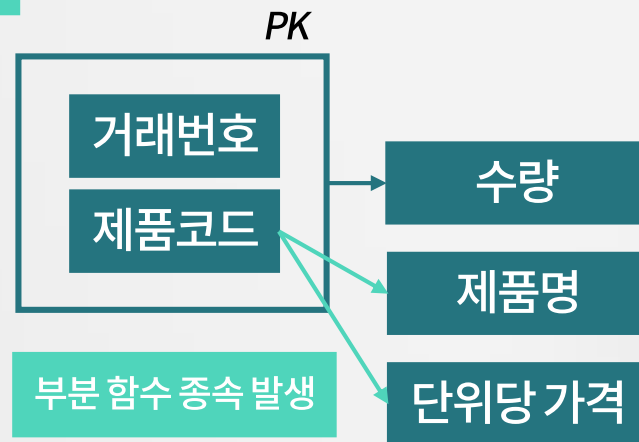
4 2차 정규화

모든 속성은 반드시 모든 기본키에 완전 함수 종속되어야 한다.
(부분 함수 종속성 제거)

PK


제2정규형 위배

거래번호	제품코드	제품명	단위당 가격	수량
5001	A01	아이스 아메리카노(R)	4,000	1
5001	A02	그린티 라떼(L)	6,000	1
5002	A01	아이스 아메리카노(R)	4,000	1
5003	A01	아이스 아메리카노(R)	4,000	2
5004	A02	그린티 라떼(L)	6,000	1

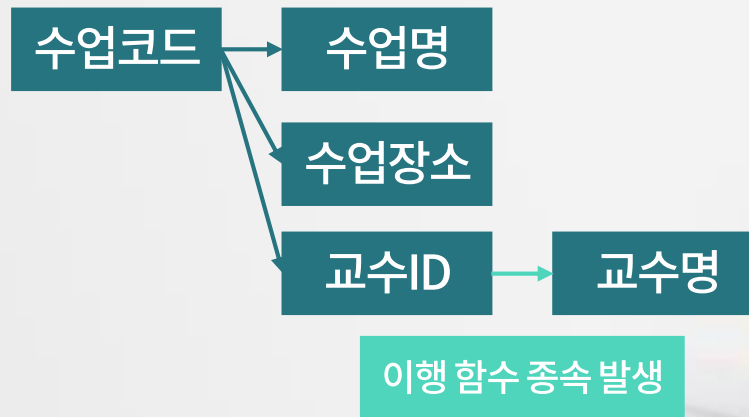


5 3차 정규화

다른 속성들이 기본키에 이행적으로 함수종속되지 않아야 한다.
(이행 함수 종속성 제거)



수업코드	수업명	수업장소	교수ID	교수명
S01	컴퓨터 활용	B-101	001	이기수
S02	경영통계학	A-205	002	김민철
S03	심리학의 이해	A-203	003	박수지



1 반정규화 개념

정규화된 엔터티, 속성, 관계에 대해 시스템의 성능향상과
개발과 운영의 단순화, 효율성을 높이기 위해
중복, 통합, 분리 등을 수행하는 데이터 모델링의 기법

- ◆ **장점**: 성능 향상과 관리의 효율성이 증가
- ◆ **단점**: 데이터 일관성, 정합성 저하
저장/유지를 위한 별도 비용이 추가적으로 발생 가능성 있음



2 반정규화 절차

① 반정규화 대상조사

- 범위처리빈도수
- 대량의 범위 처리
- 통계성 프로세스
- 테이블 조인
개수 조사

② 다른 방법유도 검토

- 뷰 생성
- 클러스터링 적용
- 인덱스 조정
- 응용어플리케이션

③ 반정규화 적용

- 테이블 반정규화
- 속성 반정규화
- 관계 반정규화



3 테이블 반정규화

테이블 병합

테이블 간의 조인(JOIN)이 많이 사용되는 경우에 이를 하나의 테이블로 병합 사용하여 성능을 높이는 방법

- ◆ 1:1 관계 테이블 병합
- ◆ 1:N 관계 테이블 병합
- ◆ 슈퍼/서브타입 관계 테이블 병합



3 테이블 반정규화

테이블 분할

◆ 수직분할

속성(칼럼) 기준으로 테이블 분할

칼럼명1	칼럼명2	칼럼명3	칼럼명4	칼럼명5

◆ 수평분할

레코드(행) 기준으로
테이블 분할

칼럼명1	칼럼명2	칼럼명3	칼럼명4	칼럼명5

3 테이블 반정규화

테이블 추가

◆ 중복테이블 추가

테이블 간의 서버가 다를 경우
동일한 테이블구조를 중복하여 생성함으로써
원격조인 제거하여 성능 향상

◆ 통계테이블 추가

집계데이터(SUM, AVG 등)를 미리 수행하여
계산해 둬으로써 조회 시 성능 향상



3 테이블 반정규화

테이블 추가

◆ 이력테이블 추가

이력에 대한 관리하기 위한 목적으로 추가하는 테이블
(이력 : 속성값들이 변경되는 경우, 변경되기 전에 대한 속성값을 저장)

◆ 부분테이블 추가

하나의 테이블의 전체 칼럼 중 자주 이용하는 특정 칼럼 있을 때,
이들만 모아 새로운 테이블 생성



4 칼럼 반정규화

◆ 중복칼럼 추가

JOIN 감소(조회 경로 단축)시키기 위해 중복된 칼럼 위치시킴

◆ 파생칼럼 추가

집계데이터(SUM, AVG 등)를 미리 수행하여 칼럼 추가

◆ 이력테이블 칼럼 추가

이력테이블에 기능성 칼럼(최근값 여부, 시작과 종료일자 등) 추가



4 칼럼 반정규화

◆ PK에 의한 칼럼 추가

복합 의미를 갖는 PK내 일부 데이터 칼럼 추가

◆ 응용시스템 오작동을 위한 칼럼 추가

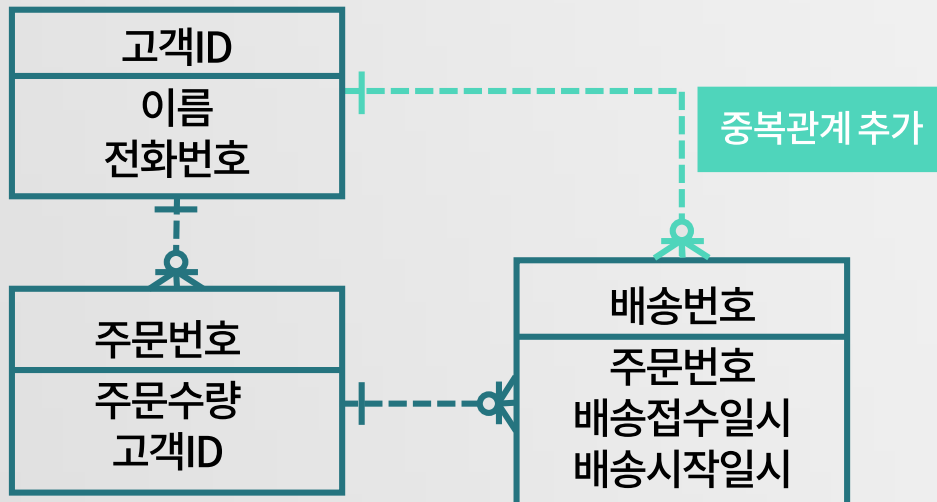
데이터 처리 시 잘못하여 이전 값으로 복구하기 원하는 경우
이전 데이터를 임시적으로 중복하여 저장하는 기법



5 관계 반정규화

◆ 중복관계 추가

데이터 처리 시, 해당 데이터 조회가 가능하지만
여러 경로를 통해야 하기 때문에 (여러 조인이 발생하기 때문에)
이 때 발생할 수 있는 성능저하 예방하기 위해 추가적인 관계 맺는 방법



1 오늘의 학습 요약

- ◆ **성능 데이터 모델링**이란 DB성능향상을 목적으로 여러 성능 관련된 모델링을 수행하는 것을 말한다.
- ◆ **함수적 종속성**이란 어떤 기준값에 의해 종속되는 현상이며, 정규화 수행 시 이에 근거하여 수행하게 된다.
- ◆ **정규화**란 반복적(중복적)인 데이터를 분리하는 것을 의미한다.
- ◆ **반정규화**란 정규화된 테이블에 대해 성능향상과 효율성을 위해 중복,통합,분리 등의 작업을 수행하는 것을 의미한다.



NEXT>>>

5

대용량 데이터에 따른 성능, DB구조와 분산 DB 데이터에 따른 성능