SAMSUNG SW ACADEMY FOR YOUTH

함께가요 미래로! Enabling People

DB 설계 및 반정규화

2024.10.21

적절한 데이터 타입 선택

- ✓ 데이터 유형은 일반적으로 문자 형/숫자 형/날짜 형으로 구분
- ✓ 같은 문자,숫자 안에서도 다양한 데이터 타입이 있으니, **저장 용량**을 고려해서 선택
  - 저장 데이터에 비해 과하지 않게 데이터 타입을 설정해야 용량 낭비가 적어짐
- ✓ 일반적으로 데이터를 저장하는데 문제가 없는 타입 중 **가장 작은 것을 선택**
  - 작은 데이터 타입일수록, 디스크나 메모리에 더 적은 공간을 사용하므로 더 빠름

DB 정규화 특징

- ✓ 테이블 간에 **중복된 데이터를 최소화**
  - **저장 용량**의 최적화 가능
  - 데이터의 **무결성** 유지
  - 쿼리 내의 **Join을 많이 사용**하게 되어, 응답시간 저하를 초래할 수 있음

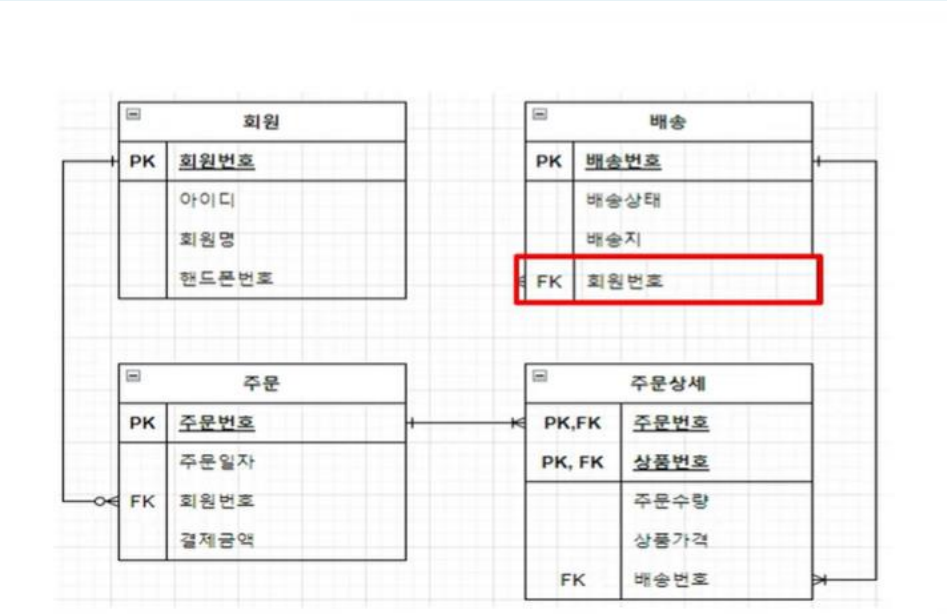
SAMSUNG SW ACADEMY FOR YOUTH

DB 반 정규화란?

- ✓ 시스템의 성능 향상, 개발과정의 편의성, 운영의 단순화를 위해 ..
- ✓ 정규화 원칙에 위배되는 행위를 의도적으로 수행하게 되는 일련의 과정
- ❖ 정규화 이후에 수행하는 것 (반정규화부터 수행하는 것은 X)

DB 반 정규화 고려 대상

- ✓ 대량의 범위를 자주 처리해야 하는 경우
- ✓ 특정 범위의 데이터만 자주 처리하는 경우
- ✓ 요약/집계 정보가 자주 요구되는 경우
- ✓ Join으로 인한 성능 저하가 예상될 때



SAMSUNG SW ACADEMY FOR YOUTH

반정규화

- ✓ 1:1관계의 테이블 병합
- ✓ 1:N관계의 테이블 병합
- ✓ 슈퍼/서브 타입 테이블 병합
- ✓ 수직분할(집중화된 일부 컬럼을 분리)
- ✓ 수평분할(행으로 구분하여 구간별 분리)
- ✓ 테이블 추가(중복테이블, 통계테이블, 이력테이블, 부분테이블)

Sharding과 Partitioning

- 공통점
  - 큰 데이터를 여러 서브셋으로 나누어 저장하는 기술
- 사용 이유
  - 성능향상과 확장성

SAMSUNG SW ACADEMY FOR YOUTH

Sharding

- ✓ 여러 인스턴스로 나눠서 저장
- ✓ 여러 샤드 노드로 데이터가 분산 저장
- ✓ 적절한 샤드키를 지정하여 해당 데이터가 존재하는 인스턴스를 찾을 수 있게 해야 함.
- ✓ 주의점 : 샤드키를 잘못 선정할 경우 한쪽으로 치우치게 될 수 있음.

Partitioning

- ✓ 하나의 인스턴스의 여러 테이블에 나누어 저장
- ✓ 인덱스 크기가 작아짐으로 조회시간을 줄임
- ✓ 큰 테이블을 제거하여 관리 용이

Partitioning

- Partitioning 종류
  - Horizontal Partitioning
    - 데이터의 개수를 기준으로 Partitioning
    - 동일한 스키마를 가진 데이터를 분산
    - Sharding도 Horizontal Partitioning의 한 종류로 볼 수 있음.
  - Vertical Partitioning
    - 테이블의 컬럼을 기준으로 데이터 분할
    - 정규화된 데이터를 분리하는 과정
    - 자주 사용하는 컬럼을 분리하여 성능 향상