

SQLD 준비자료

데이터 모델링의 이해

작성자 : 박해범

(1) ERD 작성 절차

1. 엔터티를 도출하고
그린다.

2. 엔터티를 배치한다.

3. 엔터티 간의 관계
를 설정한다.

4. 관계명을 서술한다.

5. 관계 참여도를 표현
한다.

6. 관계의 필수 여부를
표현한다.

핵심 엔터티를 좌상단에 배치한다. ex) 고객정보

(2) 엔터티의 종류

(2-1) 유/무형

유형 엔터티 : 업무에서 도출되며 지속적으로 사용되는 엔터티

개념 엔터티 : 유형 엔터티는 물리적 형태가 있지만, 개념 엔터티는 물리적 형태가 없다.

사건 엔터티 : 비즈니스 프로세스를 실행하면서 생성되는 엔터티이다.

(2) 엔터티의 종류

(2-2) 발생시점

기본 엔터티 : 키 엔터티라고도 한다. 다른 엔터티로부터 영향을 받지 않고 독립적으로 생성되는 엔터티

중심 엔터티 : 기본 엔터티와 행위 엔터티 중간에 있는 것
[기본엔터티로 부터 생성되고 행위 엔터티를 발생시킨다.]

행위 엔터티 : 2개이상의 엔터티로 부터 발생된다.

(3) 속성의 종류

(3-1) 분해 여부

단일 속성 : 하나의 의미로 구성된 것
ex) 이름, 회원 ID

복합 속성 : 여러 개의 의미가 있는 것으로 대표적으로는 주소

다중값 속성 : 여러 개의 값을 가질 수 있는 것으로 예를 들어 상품 리스트
다중값 속성은 엔터티로 분해된다.

(3) 속성의 종류

(3-2) 특성

기본 속성 : 비즈니스 프로세스에서 도출되는 본래의 속성
이름, 계좌번호, 주문일자 등

설계 속성 : 데이터 모델링 과정에서 발생하는 속성, 유일한 값을 부여한다.
ex) 상품코드, 지점코드 등

파생속성 : 다른 속성에 의해서 만들어지는 속성이다.
ex) 합계, 평균 등

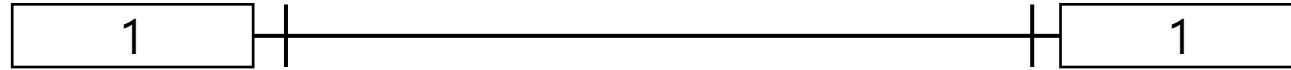
(4) 관계의 종류

존재관계 : 소속되거나 연관되는 관계 ex) 학과와 학생

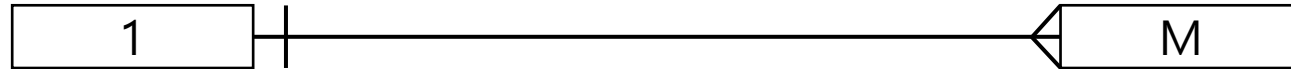
행위관계 : 관계가 행위에 의해 연결되는 관계 ex) 고객과 주문, 프로그래머와 코드

(5) 관계 차수

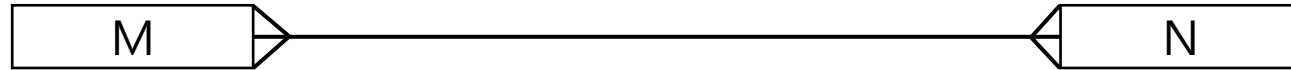
1:1 관계



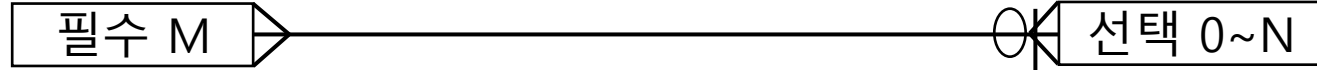
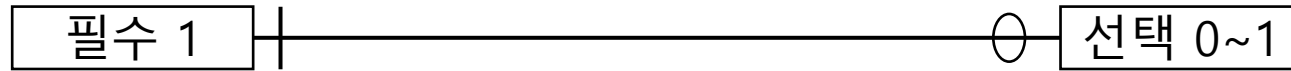
1:M 관계



M:N 관계

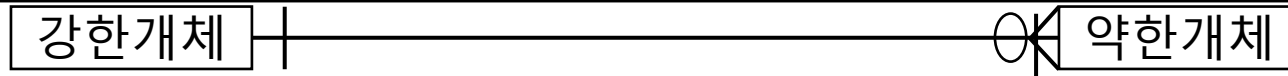


필수, 선택

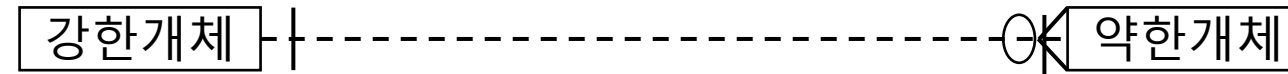


(6) 식별 / 비식별 관계

식별관계 : 강한개체의 기본키가 약한개체의 키에 포함되어 있음.(PK 공유)
강한개체가 변경되면 약한개체도 변경된다.



비식별관계 : 강한개체의 기본키가 약한개체의 칼럼에 포함되어 있음.(FK로서 연결)
점선으로 표현



(7) 식별자

pk : 기본키로 대표성, 유일성, 불변성, 최소성을 만족한다.

슈퍼키 : 최소성을 만족하지 못한다.
UNIQUE함

후보키 : 기본키가 될 수 있는 후보

외래키 : FK 다른 엔터티의 기본키를 참조한다.

대체키 : 기본키가 되지 못한 후보키

(8) 식별자 종류

(8-1) 대표성 여부

주식별자 : 유일성, 최소성을 만족하며 엔터티를 대표한다.

보조식별자 : 유일성, 최소성을 만족하지만 대표성을 만족하지 못한다.

(8) 식별자 종류

(8-2) 생성 여부

내부 식별자 : 엔터티 내부에서 스스로 생성되는 식별자.
EX) 부서코드, 주문번호, 종목코드

외부 식별자 : 다른 엔터티와의 관계로 인하여 만들어지는 식별자.
EX) 계좌 엔터티의 회원ID

(8) 식별자 종류

(8-3) 속성의 수

단일 식별자 : 하나의 속성으로 구성된 식별자.

복합 식별자 : 두 개 이상의 속성으로 구성된 식별자.

(8) 식별자 종류

(8-4) 대체 여부

본질 식별자 : 비즈니스 프로세스에서 만들어지는 식별자.

인조 식별자 : 인위적으로 만들어지는 식별자.

(8) 식별자 종류

(8-4) 대체 여부

본질 식별자 : 비즈니스 프로세스에서 만들어지는 식별자.

인조 식별자 : 인위적으로 만들어지는 식별자.

(9) 정규화

제 1 정규화 : 속성의 원자성 확보, 기본키 설정

제 2 정규화 : 부분함수종속성 제거
-> 기본키가 2개이상인 경우 분해

제 3 정규화 : 이행함수종속성 제거
-> 칼럼 간에 종속성을 제거한다.

BCNF : 후보키가 있는 경우, 후보키가 기본키를 종속 시키면 분해한다.

제 4 정규화 : 여러 칼럼들이 하나의 칼럼을 종속 시키는 경우 분해하여 다중 값 종속성을 제거.

제 5 정규화 : 조인에 의해서 종속성이 발생하는 경우 분해한다.

(10) 반정규화

(1) 계산된 칼럼추가 : 평균, 총합 등을 미리 계산해 두고 그 결과를 특정 칼럼에 추가한다.

(2) 테이블 수직분할 : 하나의 테이블을 두 개 이상의 테이블로 분할한다. 즉 칼럼을 분할하여 새로운 테이블을 만드는 것이다.

(3) 테이블 수평분할 : 하나의 테이블에 있는 값을 기준으로 테이블을 분할하는 방법이다.

cf) 파티션기법 : 값을 기준으로 DB에서 파티션을 할 수 있다.
Range partition, List partition, Hash Partition, Composite Partition

(4) 테이블 병합 : 1대1관계의 테이블을 하나로 병합하여 성능향상
1대M관계의 테이블을 병합 -> 성능은 향상되나 많은 데이터중복이 발생
슈퍼타입과 서브타입 관계가 발생하면 테이블을 통합

(10) 반정규화

(4) 테이블 병합

1대1관계의 테이블을 하나로 병합하여 성능향상

1대M관계의 테이블을 병합 -> 성능은 향상되나 많은 데이터중복이 발생

슈퍼타입과 서브타입 관계가 발생하면 테이블을 통합

(10) 반정규화

(5) 슈퍼타입 및 서브 타입 변환방법

OneToOne Type : 슈퍼타입과 서브타입을 개별 테이블로 도출한다. -> 조인이 많이 발생, 관리가 어려움

Plus Type : 슈퍼타입과 서브타입 테이블로 도출한다. -> 조인이 발생, 관리가 어려움

Single Type : 슈퍼타입과 서브타입을 하나의 테이블로 도출한다. -> 조인성능이 좋고 관리가 편하지만, 입출력 성능이 나쁘다.

(11) 분산데이터베이스

투명성

분할 투명성

위치 투명성

지역 사상 투명성

중복 투명성

장애 투명성

병행 투명성

(11) 분산데이터베이스

상향식 / 하향식 설계

상향식 설계 방식 : 지역 스키마 작성 후 향후 전역 스키마를 작성

하향식 설계 방식 : 전역 스키마 작성 후 해당 지역 사상 스키마를 작성

(11) 분산데이터베이스

장/단점

장점 :
신뢰성, 가용성 높음
병렬처리 -> 빠른 응답
용량 확장이 용이

단점 :
관리와 통제가 어렵다
보안관리가 어렵다
무결성관리가 어렵다
설계가 복잡하다