



데이터 모델링(Data Modeling)

001. 데이터 모델링의 이해



1. 데이터 모델링

- > 데이터 모델링은 현실세계를 데이터베이스로 표현하기 위해서 추상화한다.
- > 데이터 모델링을 하기 위해서는 고객과의 의사소통을 통해 고객의 업무 프로세스를 이해해야 한다.
- > 고객의 업무 프로세스를 이해한 후 데이터 모델링 표기법을 사용해서 모델링을 한다.
- > 데이터 모델링은 고객이 쉽게 이해할 수 있도록 복잡하지 않게 모델링 해야 한다.
- > 데이터 모델링은 고객의 업무 프로세스를 추상화하고, 소프트웨어를 분석/설계 하면서 점점 더 상세해진다.
- > 데이터 모델링은 고객의 비즈니스 프로세스를 이해하고 비즈니스 프로세스의 규칙을 정의한다. 정의된 비즈니스 규칙을 데이터 모델로 표현한다.



2. 데이터 모델링의 특징

- > 데이터 모델링은 추상화 해야 한다 → 추상화는 공통적인 특징을 찾고 간략하게 표현한다.
- > 데이터 모델링은 단순화해야 한다 → 복잡한 문제를 피하고 누구나 이해할 수 있게 표현한다.
- > 데이터 모델링은 명확해야 한다 → 의미적 해석이 모호하지 않고 명확하게 해석되어야 한다.

특징	설명
추상화	현실세계를 간략하게 표현한다.
단순화	누구나 쉽게 이해할 수 있도록 표현한다.
명확성	명확하게 의미가 해석되어야 하고 한 가지 의미를 가져야 한다.



3. 데이터 모델링 단계

1. 개념적 모델링

: 엔티티(Entity)와 속성(Attribute)을 도출하고 개념적 ERD(Entity Relationship Diagram)를 작성한다.

2. 논리적 모델링

: 식별자 도출, 정규화 수행

3. 물리적 모델링

: 데이터베이스를 실제로 구축한다. (ex) 테이블, 인덱스, 함수 등), 성능/보안/가용성을 고려해서 구축한다.



4. 데이터 모델링을 위한 ERD

: 엔터티와 엔터티 간의 관계를 정의하는 모델링 방법이다.

(1) ERD 작성 절차

1) 엔터티를 도출하고 그린다.

2) 엔터티를 배치한다.

3) 엔터티 간의 관계를 설정하나.

4) 관계명을 서술한다.

5) 관계 참여도를 표현한다.

6) 관계의 필수 여부를 표현한다.



5. 데이터 모델링 고려사항

1) 데이터 모델의 독립성

> 독립성을 확보하기 위해서는 중복된 데이터를 제거해야 한다.

> 데이터 중복을 제거하는 방법이 바로 정규화이다.

2) 고객 요구사항의 표현

3) 데이터 품질 확보

Quiz) 다음 중 ERD 작성 절차로 올바른 것은?

- ㄱ. 엔티티를 도출하고 그린다.
- ㄴ. 엔티티를 배치한다.
- ㄷ. 엔티티간 관계를 설정한다.
- ㄹ. 관계명을 서술한다.
- ㅁ. 관계 참여도를 표현한다.
- ㅂ. 관계의 필부 여부를 표현한다.

002. 3층 스키마 (3-Level Schema)



1. 3층 스키마

> 사용자, 설계자, 개발자가 데이터베이스를 보는 관점에 따라 데이터베이스를 기술하고 이들간의 관계를 정의한 ANSI(America National Standard Institute)표준이다.

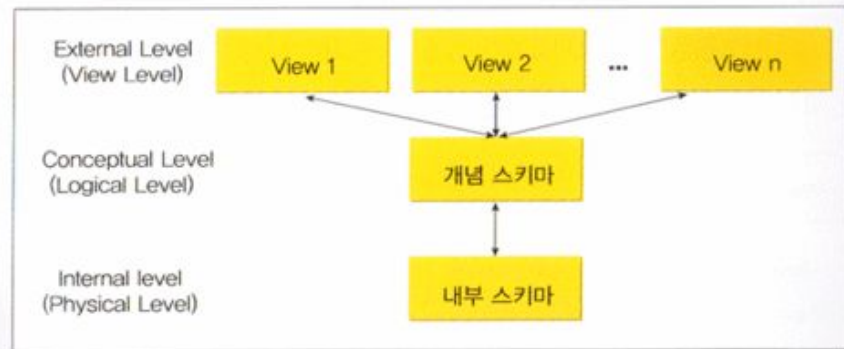
> 3층 스키마는 데이터베이스의 독립성을 확보하기 위한 방법이다.

> 데이터의 독립성을 확보하면 데이터 복잡도 증가, 데이터 중복 제거, 사용자 요구사항 변경에 따른 대응력 향상, 관리 및 유지보수 비용 절감 등의 장점을 갖는다.

> 3단계 계층으로 분리해서 독립성을 확보하는 방법으로 각 계층을 뷰(View)라고도 한다.

2. 3층 스키마 구조

[그림 2] 3층 스키마



외부 스키마: 사용자관점/ 데이터베이스뷰 (View)/응용 프로그램이 접근하는 데이터베이스

개념 스키마: 설계자관점/규칙과 구조표현/통합 데이터베이스구조

내부 스키마: 개발자관점/ 물리적 저장 구조/ 데이터의 저장, 레코드, 필드정의, 인덱스 등

Quiz) 3층 스키마의 개발자 관점에서 데이터베이스의 물리적 저장 구조를 표현하고 데이터 저장 구조, 레코드 구조, 필드 정의, 인덱스를 정의하는 것은?

ㄱ. 외부 스키마

ㄴ. 개념 스키마

ㄷ. 내부 스키마

ㄹ. 관계 스키마

003. 엔터티(Entity)



1. 엔터티(Entity)

> 엔터티는 업무에서 관리해야 하는 데이터 집합을 의미하며, 저장되고 관리되어야 하는 데이터이다.

> 엔터티는 개념, 사건, 장소 등의 명사이다.



2. 엔터티(Entity) 도출

> 엔터티는 고객의 비즈니스 프로세스에서 관리되어야 하는 정보를 추출해야 한다.



3. 엔티티(Entity) 특징

1. 업무에 필요한 정보
2. 의미있는 식별자에 의해 인스턴스는 1개씩만 존재(중복배제)
3. 2개이상의 인스턴스 집합으로 구성
4. 업무프로세스에 의해 이용되어야 함
5. 속성을 포함해야 함(식별자만 있으면 의미없음)
6. 관계가 존재해야 함

중요) 릴레이션과 테이블, 인스턴스

> 릴레이션과 테이블은 같은 의미라고 해석하면 된다. 릴레이션에 기본키 및 제약조건을 설정하면 테이블이 된다.

> 단, Relationship은 릴레이션 간의 관계를 의미한다.

> 인스턴스는 릴레이션이 가질 수 있는 **값**을 의미한다. 간단하게 생각하면 행의 수를 의미한다.



4. 엔터티(Entity)의 종류

1. 유무에 따른 분류

명칭	설명
유형 엔터티 Tangible Entity	물리적 형태가 있음 ex:사원,물품,강사
개념 엔터티 Conceptual Entity	물리적 형태가 없음 ex:조직,보험상품
사건 엔터티 Event Entity	업무 수행에 따라 발생 ex:주문,청구,미납



4. 엔터티(Entity)의 종류

2. 발생 시점에 따른 종류

명칭	설명
기본 엔터티	원래 존재하는 정보 ex:사원,부서,고객,상품
중심 엔터티	기본엔터티로부터 발생하고 다른 엔터티와의 관계를 통해 많은 행위엔터티를 발생 업무에 있어 중심 역할 ex:계약,사고,청구,주문
행위 엔터티	두 개 이상 부모 엔터티로부터 발생 내용이 자주 바뀌거나 데이터량이 증가 ex:주문목록,로그인이력

Quiz) 다음 중 발생 시점에 따른 엔터티의 종류가 아닌 것은?

ㄱ. 기본 엔터티

ㄴ. 중심 엔터티

ㄷ. 행위 엔터티

ㄹ. 사건 엔터티

004. 속성(Attribute)



1. 속성

- > 속성이라는 것은 업무에서 필요한 정보인 엔터티가 가지는 항목이다.
- > 속성은 더 이상 분리되지 않는 단위로, 업무에 필요한 데이터를 저장할 수 있다.
- > 인스턴스의 구성요소이소 의미적으로 더 이상 분해되지 않는다.



2. 속성의 특징과 종류

1) 속성의 특징

> 속성은 업무에서 관리되는 정보이다.

> 속성은 하나의 값만 가진다.

> 주식별자에게 함수적으로 종속된다. 즉, 기본키가 변경되면 속성의 값도 변경된다는 것이다.



2. 속성의 특징과 종류

2) 속성의 종류

- 특성에 따른 분류

명칭	설명
기본속성	업무로 부터 추출한 값 ex:이름,전화번호,성별
설계속성	규칙화를 위해 변형/새로정의한 값 ex:과목코드,지역코드
파생속성	다른 속성에 영향을 받아 발생한 값 ex:예금이자,평균성적



2. 속성의 특징과 종류

2) 속성의 종류

- 엔티티 구성방식에 따른 분류

종류	설명
단일 속성	- 하나의 의미로 구성된 것 ex) 이름, 회원번호 등...
복합 속성	- 여러 개의 의미가 있는 것 ex) 주소 등... - 주소는 시, 군, 구 등으로 분해가능
다중값 속성	- 속성에 여러 개의 값을 가질 수 있는 것 ex) 상품 리스트 등... - 엔티티로 분해가 된다.

Quiz) 속성중 여러 개의 값을 가질 수 있는 것으로 상품리스트와 같은 것은?

ㄱ. 단일 속성

ㄴ. 복합 속성

ㄷ. 다중값 속성

ㄹ. 결합 속성

005. 관계(Relationship)



1. 관계(Relationship)

> 관계는 엔터티 간의 관련성을 의미하며 존재 관계와 행위 관계로 분류된다.

> 존재 관계는 두 개의 엔터티가 존재 여부의 관계가 있는 것이고, 행위 관계는 두 개의 엔터티가 어떤 행위에 의한 관계성이 있는 것이다.



2. 관계의 종류

1) 존재 관계

: 존재 관계는 엔터티 간의 상태를 의미한다.

ex) 고객이 은행에 회원가입을 하면, 관리점이 할당되고, 그 할당된 관리점에서 고객을 관리한다.

2) 행위 관계

: 행위 관계는 엔터티 간에 어떤 행위가 있는 것으로, 계좌를 사용해서 주문을 발주하는 관계가 만들어진다.

ex) 증권회사는 계좌를 개설하고 주문을 발주한다.



3. 관계 차수(Cardinality)**

> 두 엔터티 간에 관계에 참여하는 수

1) 1대 1관계

종류	설명
완전 1대1(필수적)	- 하나의 엔터티에 관계되는 엔터티의 관계가 하나 있는 경우 - 반드시 존재
선택적 1대1(선택적)	- 하나의 엔터티에 관계되는 엔터티의 관계가 하나이거나 없을 수도 있다.



3. 관계 차수(Cardinality)**

2) 1대 N관계

> 1대 N 관계는 엔터티에 행이 하나 있을 때 다른 엔터티의 값이 여러 개 있는 관계이다.

ex) 고객은 여러 개의 계좌를 가질 수 있다.



3. 관계 차수(Cardinality)**

3) M대 N관계

> 두 개의 엔터티가 서로 여러 개의 관계를 가지고 있는 것

ex) 한 명의 학생이 여러개의 과목을 수강할 수 있다. 반대로 한 개의 과목은 여러 명의 학생이 수강한다.

> 관계형 데이터베이스에서 M대N 관계의 조인은 카데이안 곱이 발생한다. 그래서 M대 N관계를 1대N, N대 1로 해소해야 한다.

Quiz) 관계의 종류 중 엔터티 간의 상태를 나타내는 것은?

ㄱ. 존재 관계

ㄴ. 행위 관계

ㄷ. 속성 관계

ㄹ. 필수 관계



4. 식별 관계와 비식별 관계

1) 식별 관계

: 고객과 계좌 엔터티에서 고객은 독립적으로 존재할 수 있는 강한 개체이다.

: 강한 개체는 어떤 다른 엔터티에게 의존하지 않고 독립적으로 존재한다.

: 강한 개체는 다른 엔터티와 관계를 가질 때 다른 엔터티에게 기본키를 공유한다.

: 강한 개체는 식별 관계로 표현된다.

: 즉, 식별 관계란 고객 엔터티의 기본키인 회원ID를 계좌 엔터티의 기본키의 하나로 공유하는 것이다.

: 강한 개체의 기본키 값이 변경되면 식별관계(기본키를 공유받은)에 있는 엔터티의 값도 변경된다.



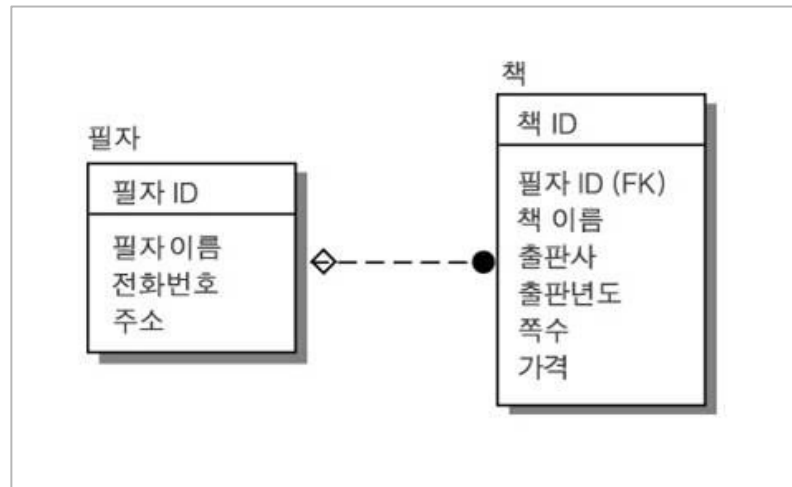
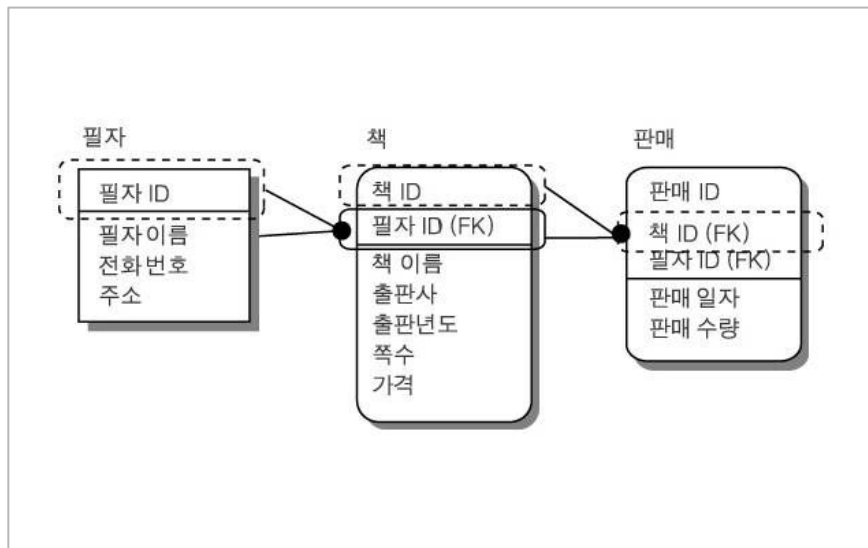
4. 식별 관계와 비식별 관계

2) 비식별 관계

: 비식별 관계는 강한 개체의 기본키를 다른 엔터티의 기본키가 아닌 일반 칼럼으로 관계를 가지는 것이다.

: 예를 들어 관리점 엔터티의 기본키는 지점 코드이고, 고객 엔터티와 비식별 관계를 가지고 있다.

즉, 지점 코드는 고객 엔터티의 기본키가 아닌 일반 칼럼으로 참조된다. 비식별 관계는 점선으로 표현한다.



[식별관계와(원) 비식별 관계(오)]

006. 엔터티 식별자(Entity Identifier)



1. 주식별자(기본키, Primary Key)

특징	내용	비고
유일성	주식별자에 의해 엔티티 내에 모든 인스턴스들을 유일하고 구분함	예) 사원번호가 주식별자가 모든 직원들에 대해 개인별로 고유하게 부여됨
최소성	주식별자를 구성하는 속성의 수는 유일성을 만족하는 최소의 수가 되어야 함	예) 사원번호만으로도 고유한 구조인데 사원분류코드 + 사원번호로 식별자가 구성될 경우 부적절한 주식별자 구조임
불변성	주식별자가 한 번 특정 엔티티에 지정되면 그 식별자의 값은 변하지 않아야 함	예) 사원번호의 값이 변한다는 의미는 이전기록이 말소되고 새로운 기록이 발생하는 개념임
존재성	주식별자가 지정되면 반드시 데이터 값이 존재해야 함 (Null 값 안됨)	예) 사원번호 없는 회사직원은 있을 수 없음



[키의 종류]

데이터베이스키	설명
기본키	후보키 중에서 엔터티를 대표할 수 있는 키이다.
후보키	후보키는 유일성과 최소성을 만족하는 키이다.
슈퍼키	슈퍼키는 유일성은 만족하지만 최소성을 만족하지 않는 키이다.
대체키	대체키는 여러 개의 후보키 중에서 기본키를 선정하고 남은 키이다.
외래키	하나 혹은 다수의 다른 테이블의 기본키 필드를 가리키는 것으로 참조 무결성을 확인하기 위해서 사용되는 키이다. 즉, 허용된 데이터 값만 데이터베이스에 저장하기 위해서 사용된다.

Quiz) 다음 중에서 참조 무결성과 관련된 데이터베이스 키는 무엇인가?

ㄱ. 기본키

ㄴ. 슈퍼키

ㄷ. 외래키

ㄹ. 대체키

2. 식별자의 종류

[표 1-1-8] 식별자의 분류체계

분류	식별자	설명
대표성 여부	주식별자	엔터티 내에서 각 어커런스를 구분할 수 있는 구분자이며, 타 엔터티와 참조관계를 연결할 수 있는 식별자
	보조식별자	엔터티 내에서 각 어커런스를 구분할 수 있는 구분자이나 대표성을 가지지 못해 참조관계 연결을 못함
스스로 생성여부	내부식별자	엔터티 내부에서 스스로 만들어지는 식별자
	외부식별자	타 엔터티와의 관계를 통해 타 엔터티로부터 받아오는 식별자
속성의 수	단일식별자	하나의 속성으로 구성된 식별자
	복합식별자	둘 이상의 속성으로 구성된 식별자
대체 여부	본질식별자	업무에 의해 만들어지는 식별자
	인조식별자	업무적으로 만들어지지는 않지만 원조식별자가 복잡한 구성을 가지고 있기 때문에 인위적으로 만든 식별자

Quiz) 다음 중 데이터베이스 식별자에 대한 종류로 올바르지 않은 것은?

- ㄱ. 식별자는 생성 여부에 따라서 내부 식별자와 외부 식별자가 있다.
- ㄴ. 식별자의 속성 수에 따라서 단일 식별자와 다중값 식별자가 있다.
- ㄷ. 식별자의 대체 여부에 따라서 본질 식별자와 인조 식별자가 있다.
- ㄹ. 식별자의 대표성 여부에 따라 주식별자와 보조 식별자가 있다.

감사합니다.

