

01

데이터베이스 개요

02

데이터베이스 시스템

03

데이터베이스 모델

학습목표

- 데이터, 정보, 데이터베이스 의미와 필요성을 알아본다.
- 정형, 비정형, 반정형 데이터의 차이점을 살펴본다.
- 데이터베이스를 사용해야 하는 이유를 알아본다.
- 데이터 독립성을 위한 3단계 데이터베이스 구조를 살펴본다.
- 다양한 데이터베이스 모델을 알아본다.
- 회원가입 화면을 이용하여 스키마 구조를 유추한다.

■ 로그인 웹 페이지

- 데이터베이스는 대용량의 데이터를 쌓아 놓은 것
- 아이디와 패스워드는 데이터베이스에 저장되어 있기 때문에 웹 사이트에서 로그인 가능

그림 9-1 로그인 웹 페이지

■ 일상생활의 데이터

- 데이터는 현실 세계에서 얻은 사실이나 측정값을 의미
- 슈퍼마켓의 물건을 보면 <물건 이름>, <가격>, <용량>, <단가>, <재고량>, <하루 판매량>, <이익> 등 모든 것이 데이터
- 데이터 유용성 여부는 사용자나 상황에 따라 변할 수 있음

■ 데이터 관리

- 데이터가 많아지면 이를 관리할 수 있는 **시스템**이 필요
- 컴퓨터를 사용하면 일일 판매 기록이나 재고량을 파악하기 쉬움
- 컴퓨터로 데이터를 가공하면 정보(information)가 됨



그림 9-2 장부에 기록하는 방식(왼쪽)과 컴퓨터로 관리하는 방식(오른쪽)

■ 데이터베이스의 정의

- 여러 사용자가 공유하고 통합해서 운영하는 데이터의 집합



■ POS 시스템(Point Of sale, POS)

- 판매와 관련한 데이터를 관리하고 정보를 수집하여 부가가치를 높이는 시스템
- 수작업으로 하던 모든 계산이나 재고 관리가 자동화 됨



그림 9-3 슈퍼마켓 계산대 앞 진열대(왼쪽)와 POS 시스템(오른쪽)

■ 정보 시스템(information system)

- 데이터베이스를 이용하여 정보 처리를 하는 시스템
- 정보 시스템으로 가공된 정보는 의사 결정에 사용

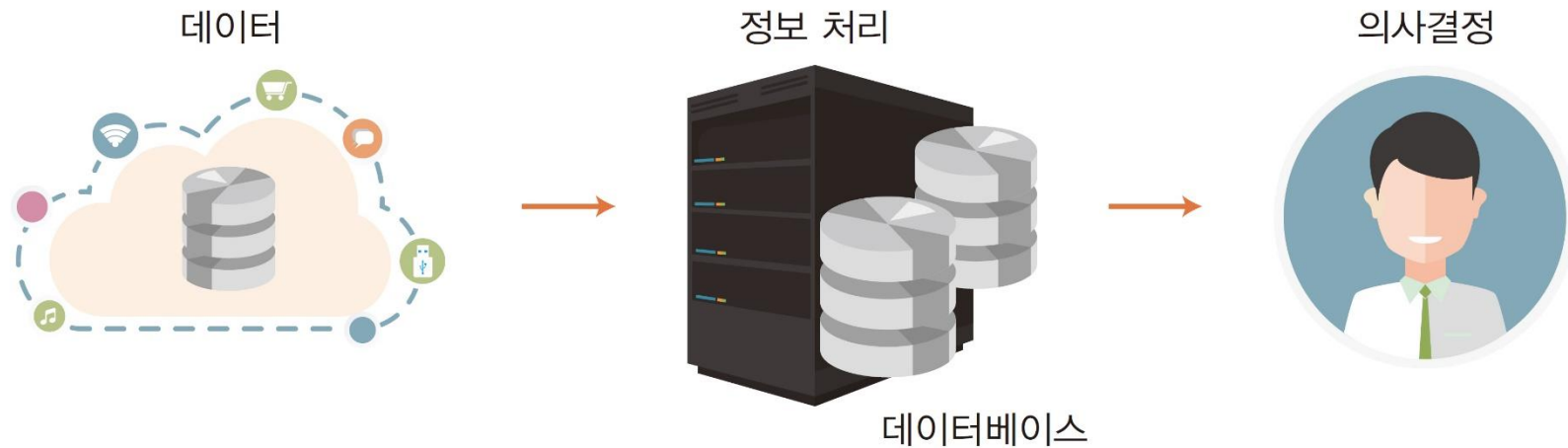


그림 9-4 정보 시스템의 구성

■ 경영 정보 시스템(Management Information System, MIS)

- 기업의 경영 관리에 필요한 정보를 수집하고 가공함으로써
기업의 생산성과 수익성을 높이는 정보 시스템

■ 의사 결정 지원 시스템(Decision Support System, DSS)

- 경영층의 의사 결정자에게 정보를 분석하고 도식화하여 제공함으로써
효과적으로 의사 결정 과정을 진행할 수 있게 하는 정보 시스템

■ 콘텐츠 관리 시스템(Content Management System, CMS)

- 문서, 이미지, 동영상 등 다양한 종류의 콘텐츠를 생성, 보관, 관리하는 시스템
- 다양한 문서 관리, 이력도 관리하여 외부 유출이나 불법적인 복사를 막아줌

■ 교육 관리 시스템(Learning Management System)

- 교육 진행 시 각 학생의 학습 진도 관리뿐 아니라 교육 콘텐츠 생성 및 활용에도 사용하는 시스템

■ 데이터 마이닝(data mining)

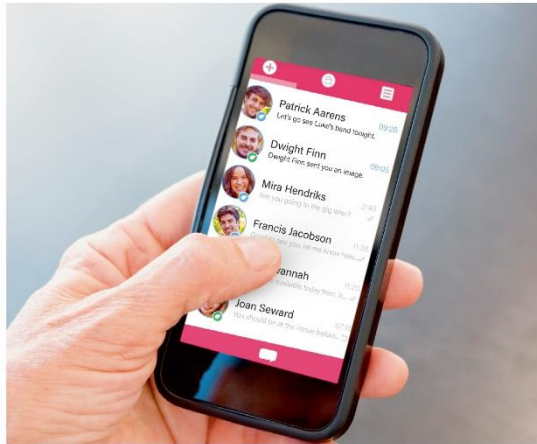
- 데이터에서 유용한 정보를 추출하는 기술

■ 빅데이터(big data) 기술

- 데이터 마이닝 기술을 이용하여 엄청나게 큰 데이터에서 유용한 정보를 추출하는 것

■ 정형 데이터(structured data)

- 미리 정해진 구조에 따라 저장된 데이터
- 관리가 용이

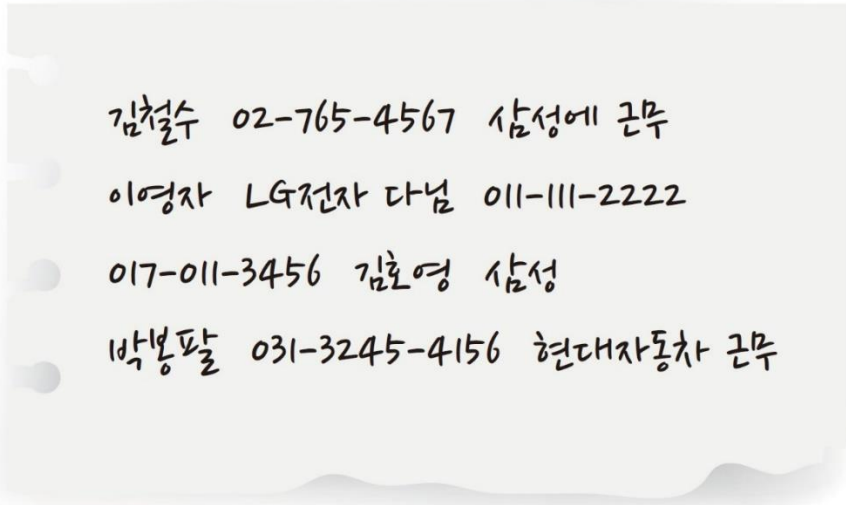


이름	전화번호	직장
김철수	02-765-4567	삼성
이영자	011-111-2222	LG전자
김호영	017-011-3456	삼성
박봉팔	031-3245-4156	현대자동차

그림 9-5 정형 데이터의 예 - 스마트폰 전화번호부

■ 비정형 데이터(unstructured data)

- 정해진 구조가 없는 데이터
- 일반적으로 트위터나 페이스북의 SNS데이터, 언어, 음악, 사진과 같이 특별한 형식이 없이 만들어진 데이터



김철수 02-765-4567 삼성에 근무
이영자 LG전자 다닐 011-111-2222
017-011-3456 김호영 삼성
박병팔 031-3245-4156 현대자동차 근무

그림 9-6 비정형 데이터 예 - 수첩에 적는 전화번호부

■ 반정형 데이터(semi-structured data)

- 정형과 비정형의 중간 형태인 데이터
- 구조와 데이터를 같이 저장
- 구조 정보를 나타내는 태그(tag) 사용
- 예: XML

〈이름〉김철수〈/이름〉

〈전화〉02-765-4567〈/전화〉

〈직장〉삼성〈/직장〉

〈이름〉이영자〈/이름〉

〈직장〉G전자〈/직장〉

〈전화〉011-111-2222〈/전화〉

그림 9-7 반정형 데이터 예 - XML

■ 데이터베이스의 특징

- 실시간 접근 : 실시간으로 데이터에 접근 가능
- 최신 데이터 유지 : 데이터의 삽입, 삭제, 수정으로 항상 최신 데이터 유지
- 동시 공유 : 여러 사용자가 동시에 데이터 공유 가능
- 내용 참조 : 내용을 참조하여 데이터에 접근 가능

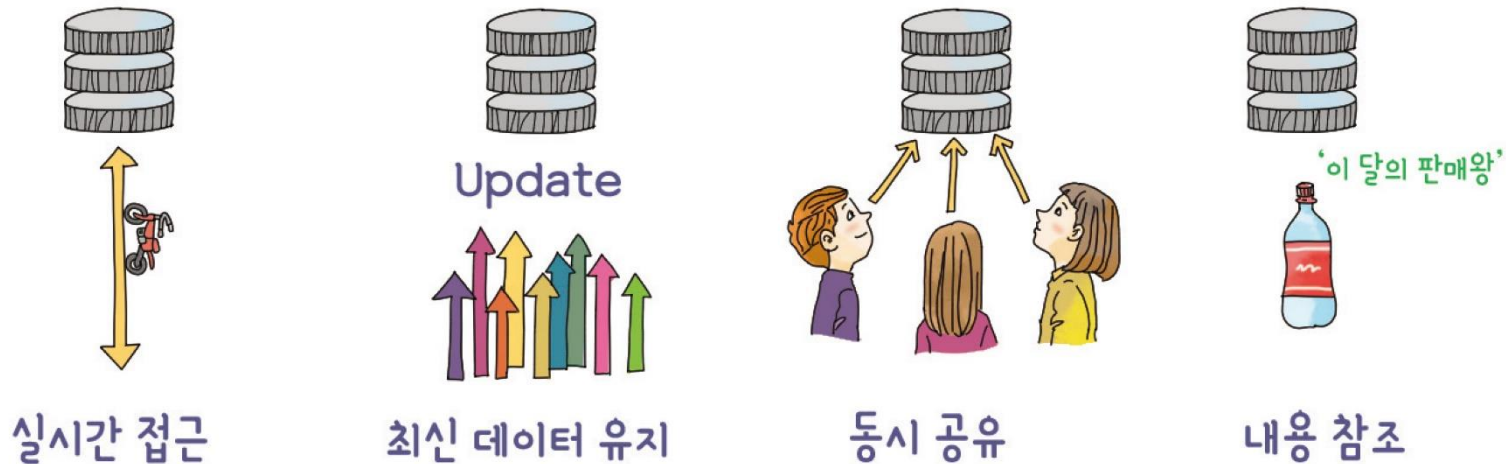


그림 9-8 데이터베이스의 특징

■ 데이터 중복성

- 같은 데이터가 여러 프로그램에 중복되어 나타나는 문제
- [예] 콜라 가격을 1,000원에서 2,000원으로 변경할 경우
 - 판매 프로그램의 판매가 변경
 - 재고 프로그램의 판매가 변경
 - 결국 사용하는 모든 프로그램을 변경
- 여러 프로그램이 같은 데이터파일을 공유하면 중복성을 해소 할 수 있지만, 종속의 문제 발생



그림 9-9 데이터 중복성의 문제

■ 데이터 종속성

- 여러 프로그램이 같은 데이터 파일을 공유하는 경우
=> 프로그램이 데이터 변화에 종속되는 문제 발생
- [예] 판매 프로그램에 '할인율'이라는 새로운 데이터 추가할 경우
 - 판매 프로그램의 '재고' 앞에 '할인율'이라는 새로운 칸 생성
 - 새로운 파일 구조에 맞도록 재고 프로그램, 납품 프로그램도 변경
 - 결국 연쇄적으로 파일을 공유하는 모든 프로그램 변경



그림 9-10 데이터 종속성의 문제

■ 보안 문제

- 허가받은 응용 프로그램이나 사용자만 데이터에 접근할 수 있어야 하며, 그러한 보안책 필요



그림 9-11 보안의 문제

■ 백업 및 회복의 문제

- 실수로 데이터를 지운 경우, 백업이 없으면 데이터가 사라짐
- 누군가에게 백업을 지시했다고 할지라도 누가 얼마큼 자주 백업을 할지 결정하는 것은 쉽지 않음



그림 9-12 백업 및 회복의 문제

■ 데이터베이스 관리 시스템(DataBase Management System, DBMS)

- 데이터베이스를 관리하고 작업을 통해 의미 있는 정보를 생성할 수 있는 시스템
- 데이터베이스에 쌓인 데이터의 삽입, 삭제, 검색 등을 쉽게 할 수 있게 해줌
- 파일을 사용할 때 발생하는 데이터 중복성, 종속성, 보안, 백업과 회복의 문제 모두 해결 가능
- 자신을 통해서만 데이터베이스에 접근 가능하므로 데이터베이스를 보호함
- DBMS로 인해 사용자는 데이터가 어떻게 저장되었는지 알 수도 없고, 알 필요도 없음

■ SQL(Structured Query Language)

- 데이터베이스에 저장된 데이터를 처리하려고 만든 언어
- SQL은 스키마를 만들어 관리하는데, 스키마는 데이터의 삽입, 삭제, 갱신 등의 작업에 적합한 구조를 지님
- 질문(query)을 하면 DBMS가 응답을 하는 구조이기 때문에 Structured Query Language라고 함

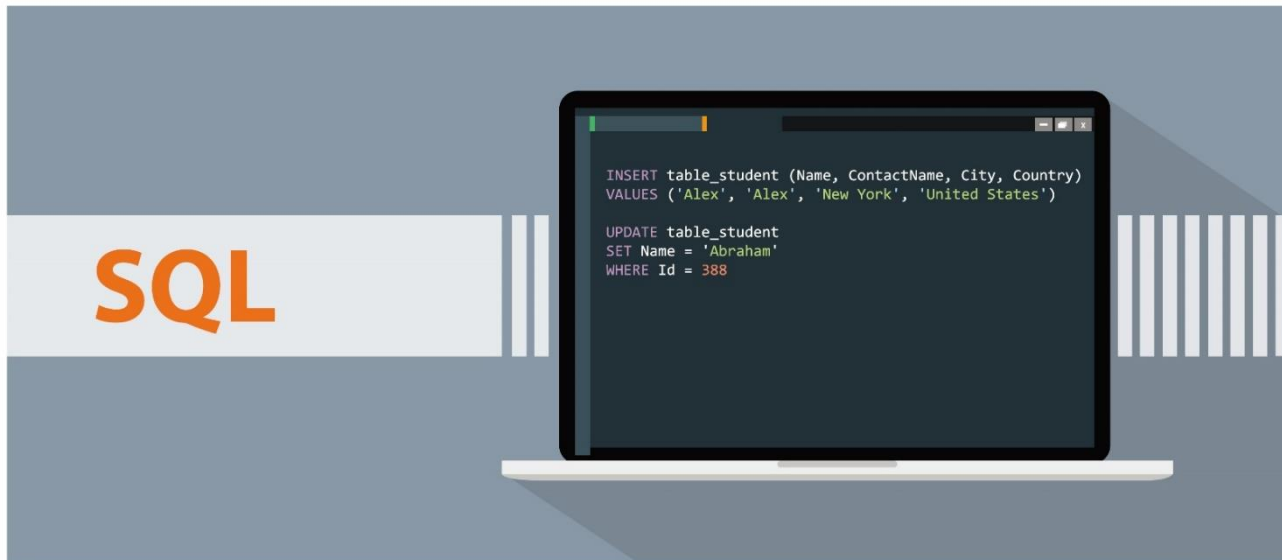


그림 9-13 SQL(Structured Query Language)

■ 데이터베이스 시스템

- 데이터를 저장한 ‘데이터베이스’와 데이터베이스를 관리하는 시스템인 ‘DBMS’을 합쳐서 부르는 말
- 프로그램이나 사용자가 입력한 SQL을 해석하여 응답하고, 데이터베이스에 저장된 데이터가 훼손되지 않도록 관리하고 백업하는 역할을 함

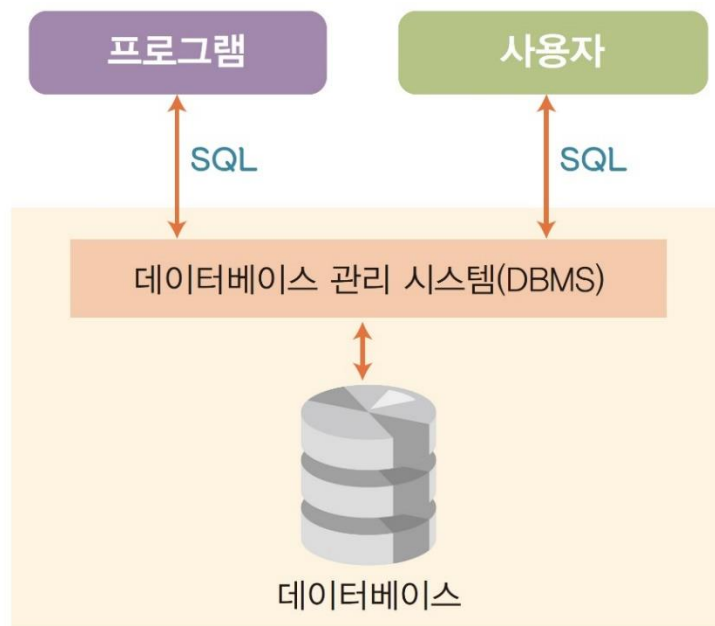
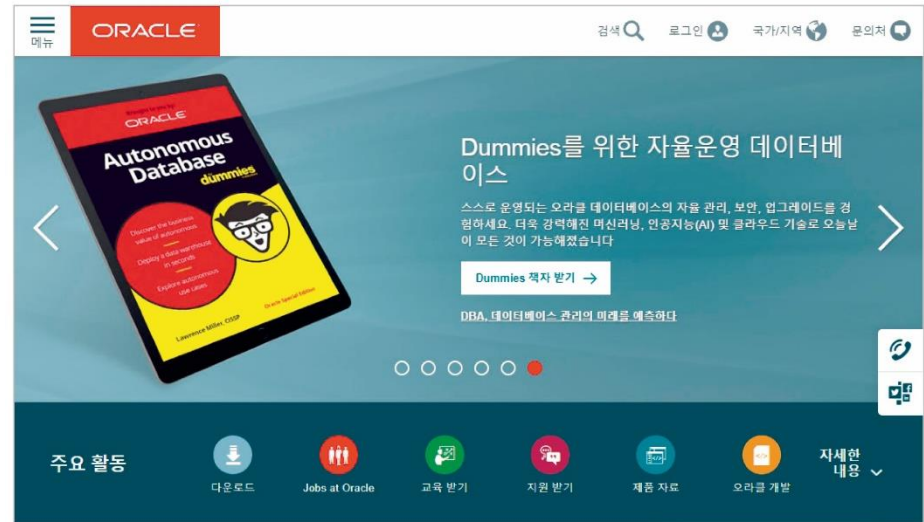


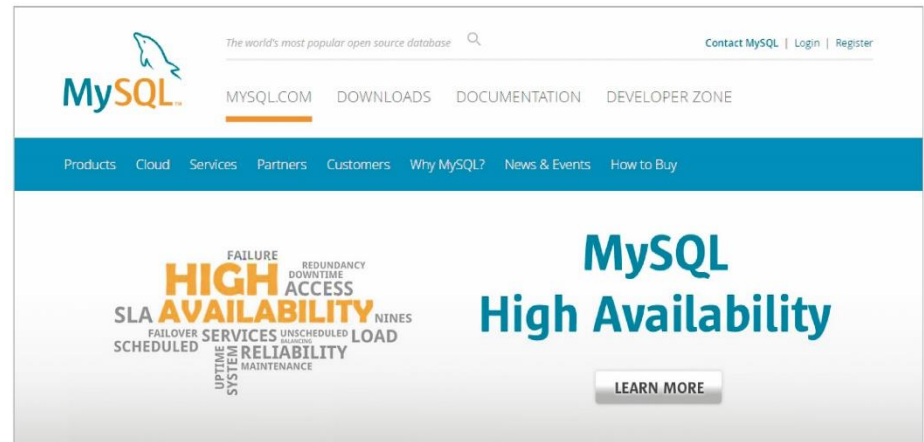
그림 9-14 데이터베이스 시스템의 구조

■ DBMS의 종류

- 오라클 : 상업용 DBMS로 가장 성공한 제품
- MySQL : 무료로 사용 가능한 DBMS



(a) 오라클 홈페이지



(b) MySQL 홈페이지

그림 9-15 DBMS 제품 홈페이지 화면

■ 스키마(schema)의 정의

- 스키마는 데이터베이스에 저장되는 데이터 구조와 제약 조건을 정의한 것
- 일종의 데이터베이스 설계도면 같은 것으로, 처음부터 잘 만들지 않으면 문제가 발생
- 하나의 데이터를 저장할 수 있는 영역이 필드
- 하나의 스키마는 연관 있는 필드들의 집합

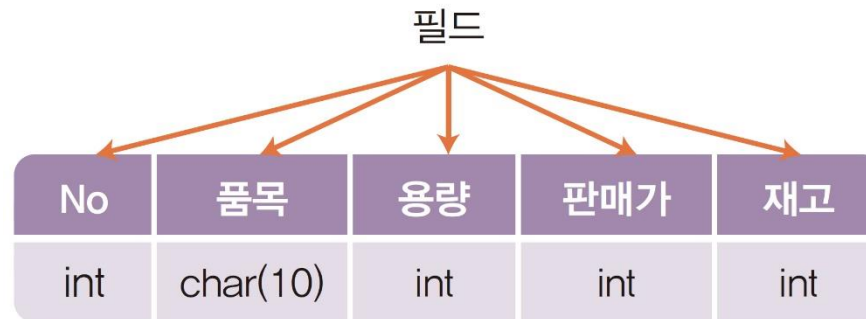


그림 9-16 스키마의 예

■ 레코드

- 스키마에는 각 필드의 제약 사항이 명시되어 있음
 - 품목은 문자 열 글자, 용량은 정수(int)처럼 데이터의 형태 및 크기 명시
- 스키마 구조에 맞추어 저장된 데이터 묶음 1개를 **레코드(record)**라고 함
- 레코드 여러 개가 채워지면 하나의 파일이 됨
- 데이터베이스는 관련 있는 파일을 묶어 하나의 데이터 묶음인 것처럼 관리



그림 9-17 레코드의 예

■ 3단계 데이터베이스 구조(3-level-database architecture)

- 3단계로 추상화된 데이터베이스 구조
- 데이터베이스를 사용자, 조직, 시스템 관점에 따라 외부 단계, 개념 단계, 내부 단계로 나눔
- 미국표준화기관인 ANSI/SPARC에서 제안

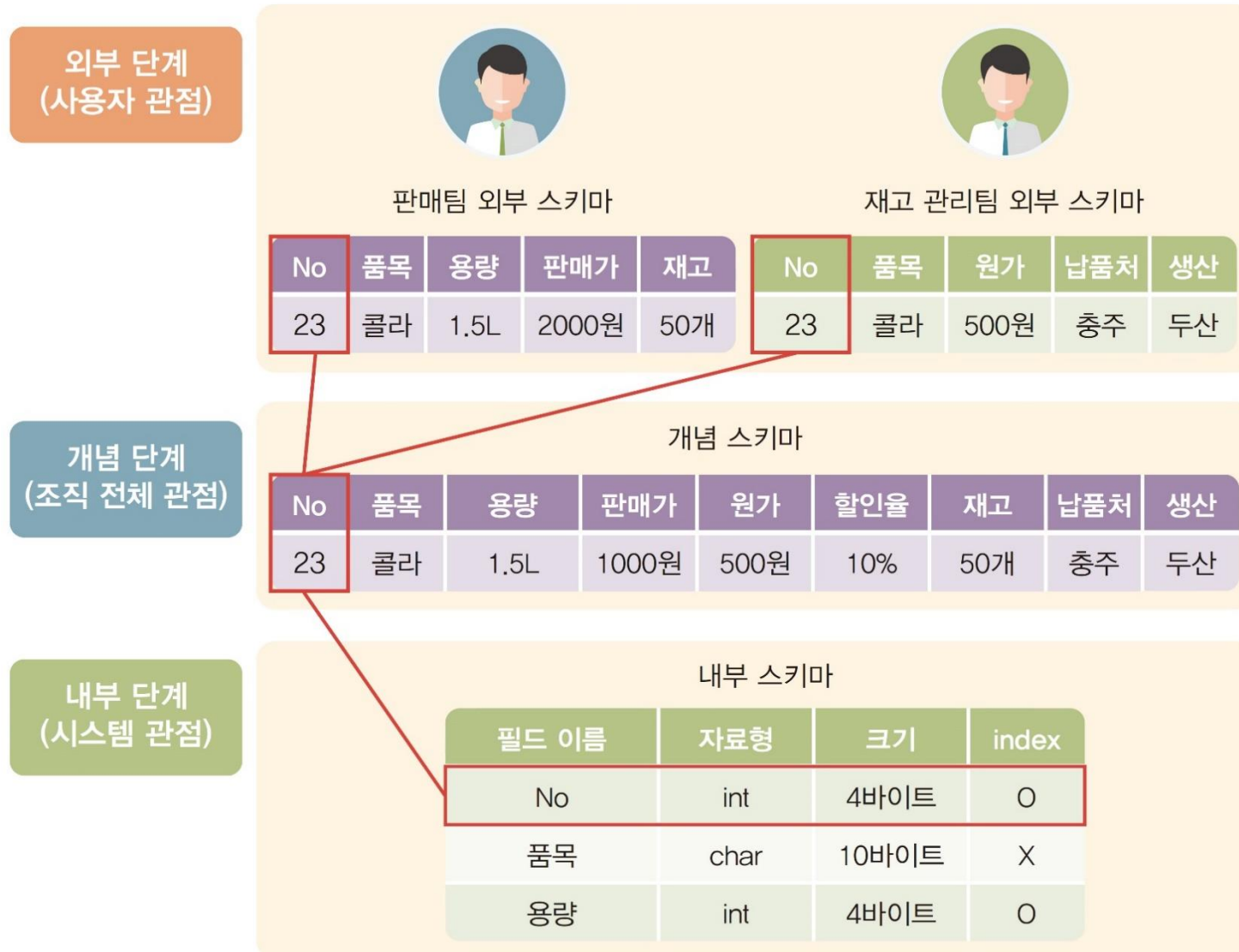


그림 9-18 3단계 데이터베이스 구조

■ 외부 단계(external level)

- 사용자 관점에서 보이는 데이터베이스
- 사용자 또는 프로그램에 보이는 단계이므로 뷰 단계(view level)라고도 함
- 전체 데이터베이스를 모든 사람이 볼 수 있게 하는 것은 위험하여
전체 스키마 중에서 각 직원이나 팀에 필요하다고 생각되는 스키마를
새로 구성하여 보여 줌
- 외부 스키마(external schema) 또는 서브 스키마(sub-schema)라고 함
- 전체 스키마에서 외부 스키마를 여러 개 생성 가능

■ 개념 단계(concept level)

- 조직 전체의 관점에서 바라보는 데이터베이스
- 전체 데이터베이스를 구현하는 단계로 모든 외부 스키마를 통합하여 하나의 스키마 즉, 개념 스키마(concept schema)로 구성한 것
- 개념 스키마는 전체 시스템에서 사용할 업무와 권한이 개념적으로 정해짐
- 개념 스키마는 데이터베이스 내에 1개만 존재
- 개념 스키마에는 각 데이터 간의 관계나 제약 사항, 접근 권한, 보안 정책 등이 정의되어 있음
- 개념 스키마에는 데이터를 물리적으로 어떻게 저장할지를 정하는 정보는 없기 때문에 일반적으로 스키마라고 하면 개념 스키마를 가리킴

■ 내부 단계(internal level)

- 데이터베이스를 실제로 저장하는 단계
- 디스크나 테이프 등 저장 장치에 어떤 방식으로 저장할지 결정
- 레코드 구조, 필드의 종류와 크기, 접근을 할 수 있는 인덱스(index)처럼 저장에 필요한 정보들을 결정하는 단계
- 내부 스키마(internal schema)라고도 함

■ 데이터 독립성(data independency)

- 데이터베이스를 3단계로 추상화하고 단계별로 스키마를 만드는 것은 데이터의 독립성 때문
- 외부 스키마-개념 스키마-내부 스키마의 변경이 서로에게 영향을 미치지 않고 독립적으로 움직인다는 의미

외부 단계
(사용자 관점)



‘용량’과 ‘판매가’
순서를 바꾸어 주세요.

외부 스키마

No	품목	용량	판매가	재고
23	콜라	1.5L	2000원	50개

개념 단계
(조직 전체 관점)

개념 스키마

‘세일가’ 필드 추가

No	품목	용량	판매가	원가	세일가	할인율	재고	납품처	생산
23	콜라	1.5L	1000원	500원	900원	10%	50개	충주	두산

내부 단계
(시스템 관점)

내부 스키마

‘No’ 필드의 크기를
8바이트로 변경

필드 이름	자료형	크기	index
No	int	8바이트	O
품목	char	10바이트	X
용량	int	4바이트	O

그림 9-19 데이터의 독립성

■ 계층 모델

- 초기 데이터베이스
- 데이터와 데이터 관계에 계층(structure) 구조를 사용
- 구조적인 제약으로 데이터들의 관계를 표현하기가 어려움

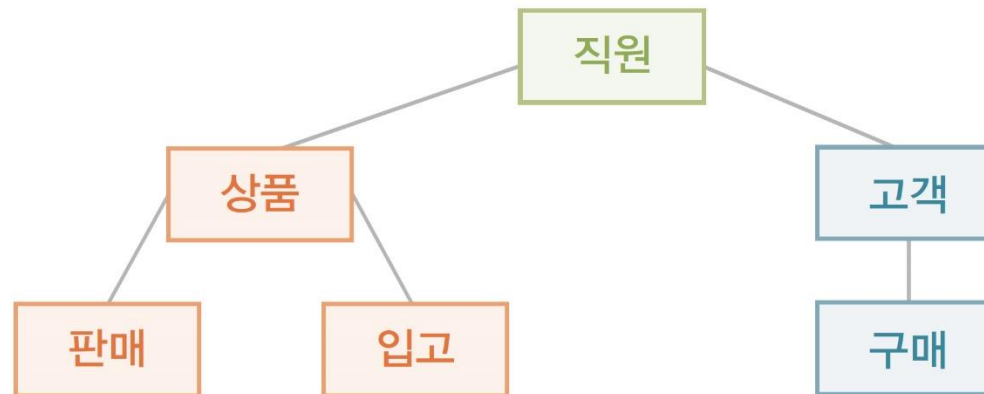


그림 9-20 계층 모델

■ 네트워크 모델

- 데이터 관계를 선 여러 개로 연결한 구조
- 데이터 관계를 표현하기에 적합하지만, 구조가 복잡하다는 단점이 있음

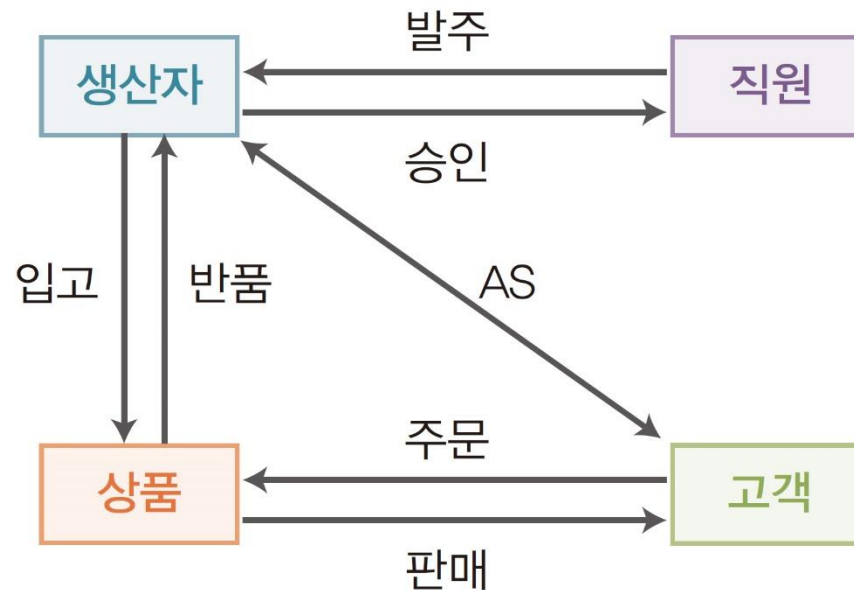


그림 9-21 네트워크 모델

■ 관계 모델

- 데이터는 테이블 형태로 구성하고, 각 테이블들은 서로 간의 키(key)를 중심으로 관계를 형성하는 구조
- 사용하기 쉽고 구조가 간단하여 대부분의 데이터베이스가 사용
- 관계 데이터베이스 관리 시스템을 RDBMS(Relation DataBase Management System)라고 함

판매 데이터베이스

No	품목	용량	판매가	재고
23	콜라	1.5L	2000원	50개

제품 데이터베이스

No	품목	용량	입고가	유통기한
23	콜라	1.5L	500원	2년

그림 9-22 관계 모델

■ 객체 모델

- 객체 언어처럼 데이터와 그 데이터를 처리하는 함수를 하나의 객체 형태로 구성한 데이터베이스
- 데이터와 그 데이터를 처리하는 데 필요한 함수를 하나로 묶어 객체로 만듦
- 객체 모델에서는 값을 입력할 때마다 문제가 있는지 자동으로 검사하기 때문에 따로 입력 값을 검사할 필요가 없음

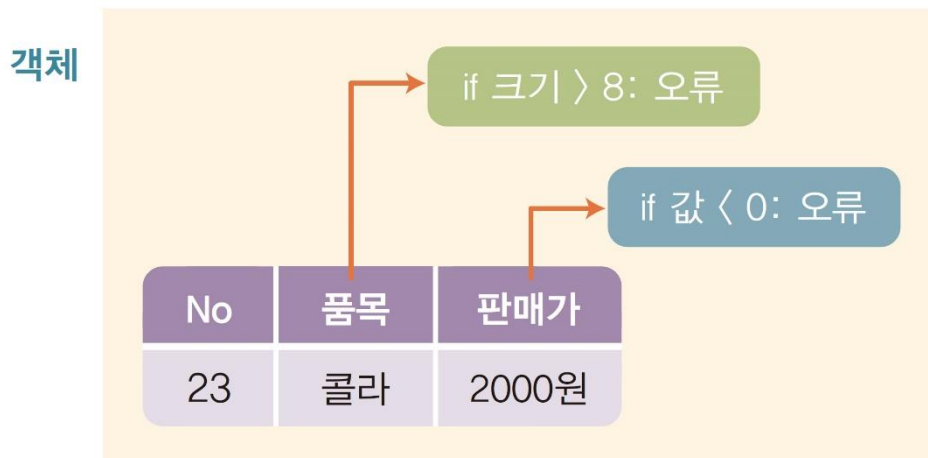


그림 9-23 객체 모델

■ NoSQL 모델

- 음악이나 사진처럼 **비정형 데이터를 처리**하려고 만든 데이터베이스
- SQL이 아닌(No) 비정형 데이터 처리에 맞는 새로운 언어 구조를 사용
- 일부 기업이나 연구소에서 사용(현재까지 RDBMS가 주류)

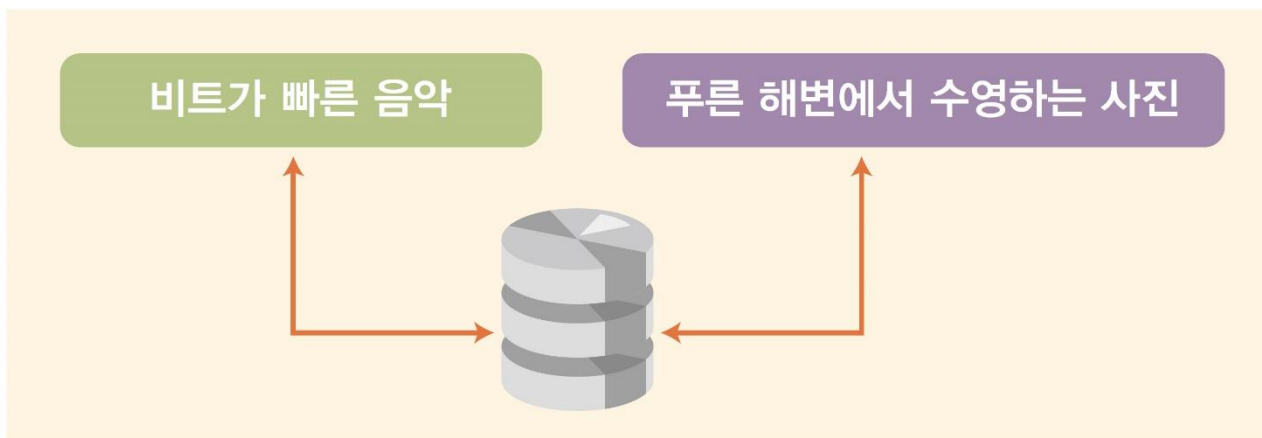


그림 9-24 NoSQL 모델

■ 데이터 모델링

- 현실의 데이터를 데이터베이스에 저장할 수 있는 데이터로 변환하는 과정
- 데이터를 모델링하여 만든 하나의 테이블을 개체(entity)라고 함

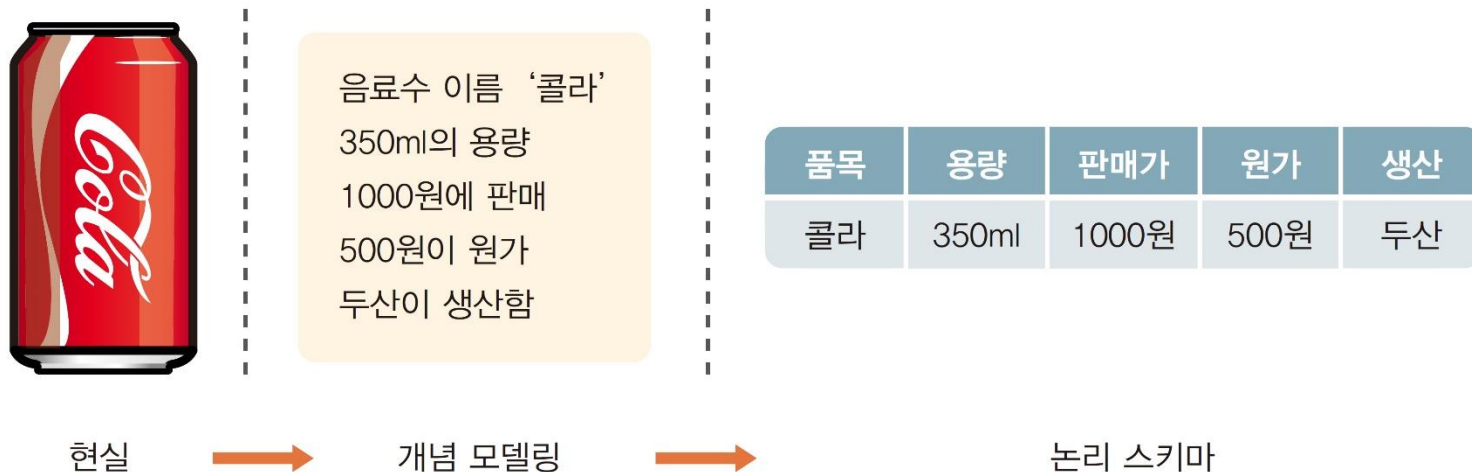


그림 9-25 콜라의 데이터 모델링 예

■ 관계 데이터베이스

- 개체와 개체의 관계(relationship)를 연결한 데이터베이스
- 구조화된 데이터와 제약 조건을 그래프 형식으로 표현하는 방법으로 개체-관계 모델링(Entity-Relationship Modeling, ERM)을 사용
- 개체-관계 모델링의 결과물을 개체-관계 다이어그램(Entity-Relationship Diagram, ERD)이라고 함

■ 개체-관계 모델링의 구성 요소

- 개체(entity) : 현실 세계를 구성하는 사람이나 사물처럼 구별이 가능한 모든 것
- 속성(attribute) : 개체가 가지는 고유의 특성
- 관계(relationship) : 개체와 개체가 맺고 있는 연관성

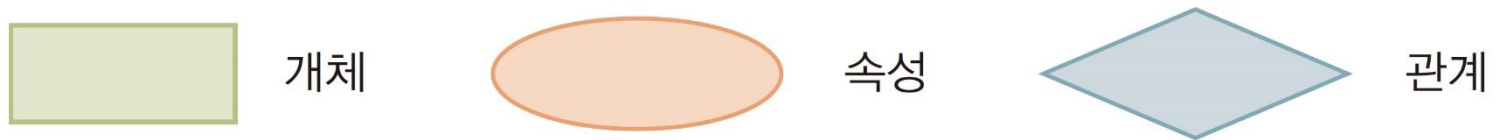


그림 9-26 개체-관계 다이어그램의 구성 요소

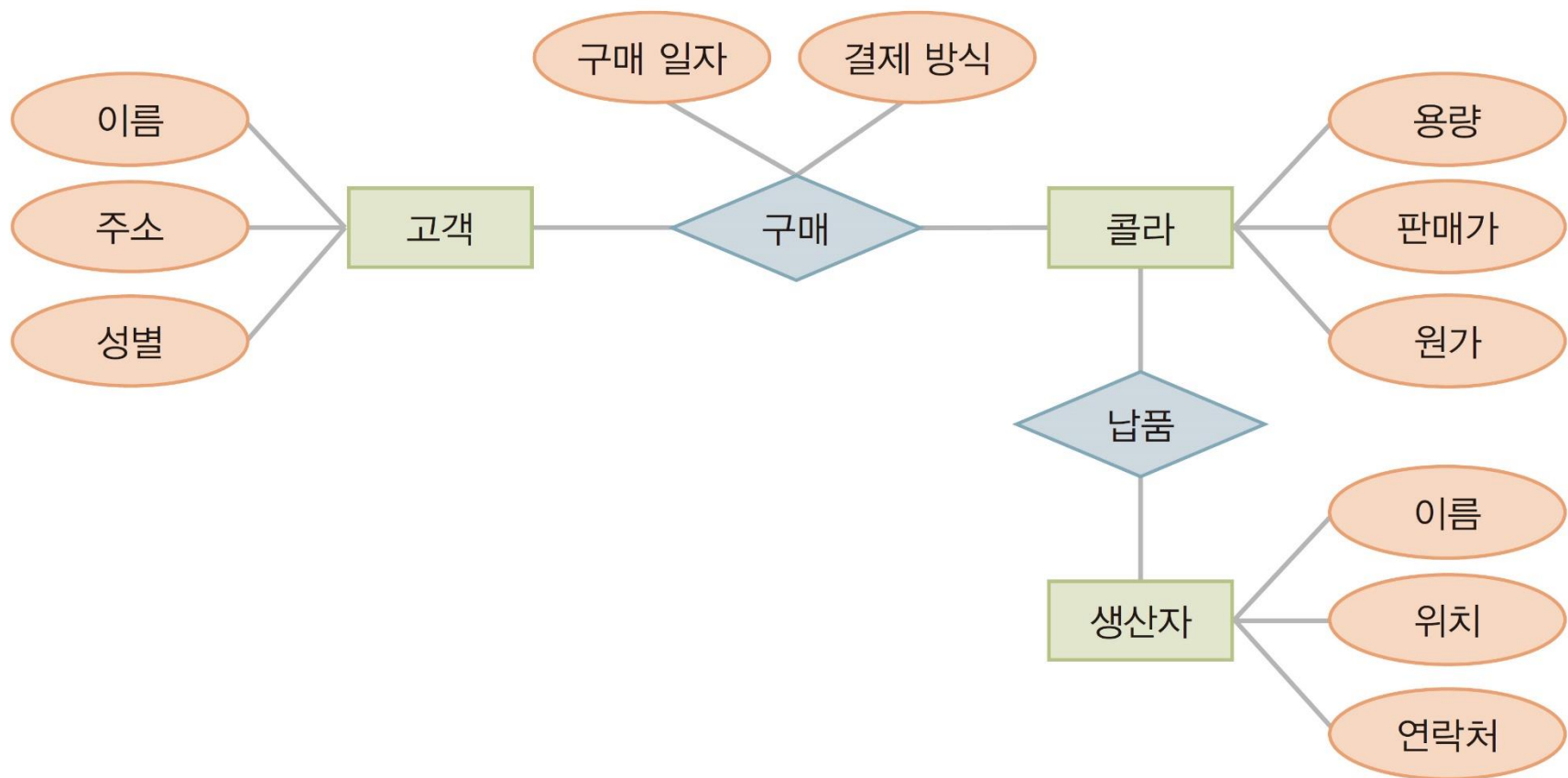


그림 9-27 개체-관계 다이어그램의 예

■ 개체 테이블

- 개체-관계 모델링이 끝나면 각 개체의 테이블 구조(스키마)를 설계함
- 테이블 구조(스키마) 설계 시 ID 필드는 여러 레코드 중 각 레코드를 접근할 수 있는 대푯값으로 구성

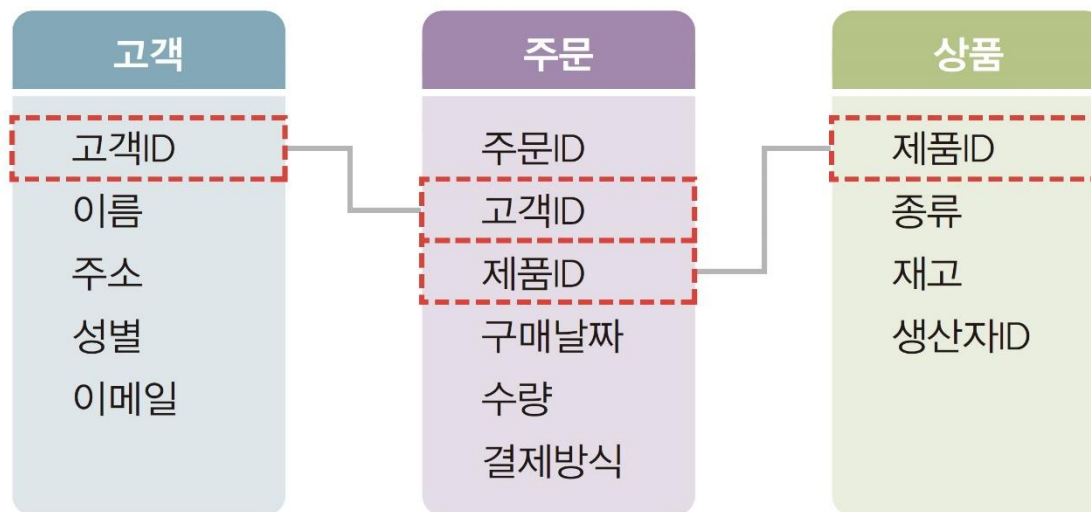


그림 9-28 개체 테이블의 예

■ 데이터베이스 설계

- 데이터베이스에서 가장 중요한 단계가 ER 모델링과 스키마 설계
- 데이터베이스는 설계를 잘못하면 처음부터 다시 시작해야 함

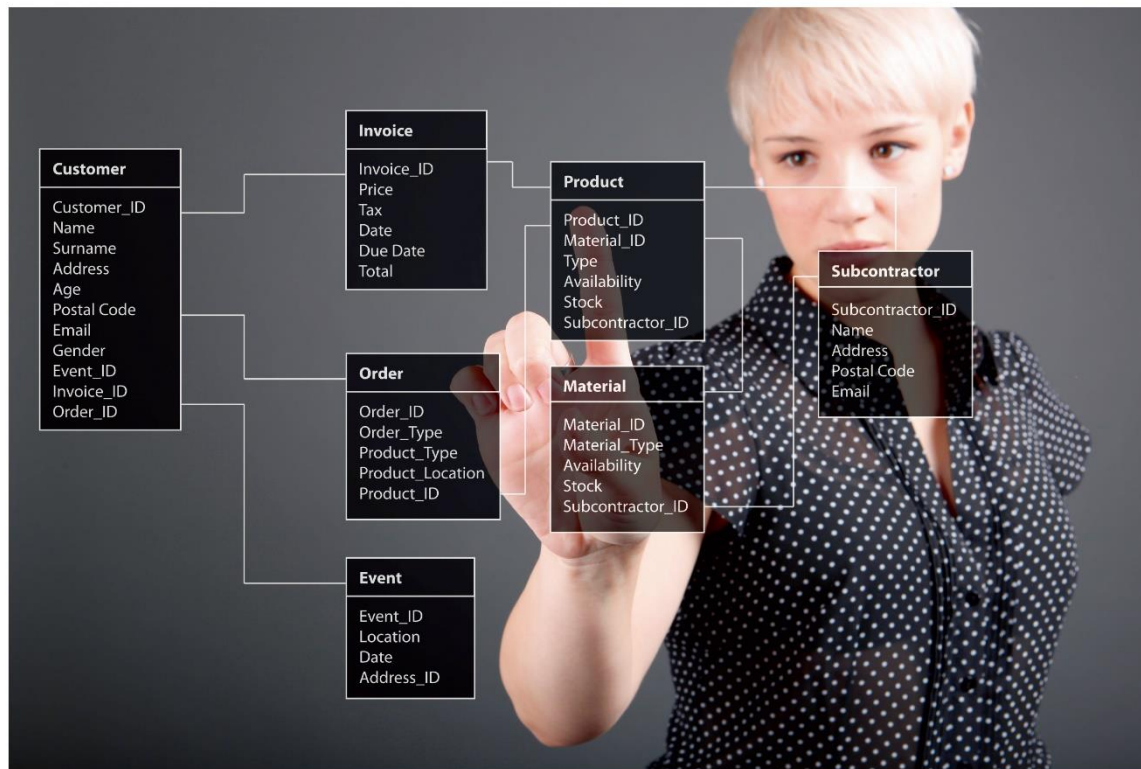


그림 9-29 데이터베이스의 설계

■ 네이버 회원가입 예

- 회원가입 화면에 나타나는 필드는 사용자의 스키마 필드 구성과 같음
- 네이버 사용자의 스키마 필드는 [아이디, 비밀번호, 이름, 생년월일, 성별, 이메일, 휴대전화]로 구성
- 회원가입에 있는 모든 필드를 다 채운 후 <회원가입> 버튼을 누르면 SQL이 작동하여 입력 내용이 사용자 스키마에 삽입됨

The image shows a web form for Naver registration. The fields are as follows:

- 아이디**: A text input field with a placeholder "@naver.com" and a lock icon on the right.
- 비밀번호**: A text input field with a lock icon on the right.
- 비밀번호 재확인**: A text input field with a lock icon on the right.
- 이름**: A text input field.
- 생년월일**: Three separate input fields for year (placeholder "년(4자)"), month (placeholder "월" with a dropdown arrow), and day (placeholder "일").
- 성별**: A dropdown menu with the placeholder "성별" and a dropdown arrow.
- 본인 확인 이메일 (선택)**: A text input field with the placeholder "선택입력".
- 휴대전화**: A dropdown menu with the placeholder "대한민국 +82" and a dropdown arrow.

그림 9-30 네이버 회원가입 화면

■ 네이버 회원가입 예

- 입력된 정보는 네이버의 사용자 데이터베이스에 있는 레코드 중 하나의 레코드가 됨

레코드로 저장

The image shows a Naver registration form with the following fields:

- 아이디 (ID): Input field with a placeholder "@naver.com".
- 비밀번호 (Password): Input field with a lock icon.
- 비밀번호 재확인 (Confirm Password): Input field with a lock icon.
- 이름 (Name): Input field.
- 생년월일 (Date of Birth): Fields for year (4 digits), month (dropdown), and day (dropdown).
- 성별 (Gender): Dropdown menu with "성별" (Gender) as the selected option.
- 본인 확인 이메일 (선택) (Verify Email (Optional)): Input field with "선택입력" (Optional input) as the placeholder.
- 휴대전화 (Mobile Phone): Input field with "대한민국 +82" (South Korea +82) as the selected country code.


그림 9-30 네이버 회원가입 화면

고객ID	아이디	비밀번호	이름	생년월일	성별	이메일	휴대전화
21248	SHCho	*****	조성호	1999 09 09	남	zoch@hs.ac.kr	017-234-21


그림 9-31 사용자 스키마와 레코드

■ 네이버 물건 구매

- 네이버에 로그인할 때마다 내 레코드의 아이디와 비밀번호를 이용하여 내가 맞는지 확인
- 이메일을 보내거나 물건을 구매할 때마다 해당 정보는 데이터베이스에 쌓임

상품주문번호	상품정보	상품금액(수량)	배송비/판매자	진행상태	
2017102510009160	 <p>[시디즈] 무료배송 T50 AIR 전용 좌판 커버(T529D) 색상:다크그레이 (무료배송)</p> <p>재구매 ▼</p>	23,000원 (1개)	무료 퍼니엄 (02-2275-3236) 판매자 문의	구매확정	배송조회

결제금액정보					현금영수증 정보확인
상품금액	할인금액	환불정산액/포인트 결제액		상품금액	23,000원
└ 상품합계 23,000원	└ 상품/스토어/주문합인 0원	└ 포인트 1,613원		할인금액	(-) 0원
└ 배송비합계 0원	└ 배송비합인 0원	1,613원		포인트	(-) 1,613원
23,000원	0원				

 Pay 혜택 (최대 330원 적립)

· 구매 적립 : 구매 확정시 230원 적립예정 ?
 · 리뷰 적립 : 리뷰 작성시 최대 100원 적립예정 ?

! 동일 상품(상품 번호 기준)에 대한 리뷰 적립 혜택은 1회로 제한되며 적립 후 30일 경과 시 리뷰 적립 혜택을 다시 받을 수 있습니다

카드 간편결제		
하나 (일시불	

배송지 정보		주문자 정보
수령인	조성호	회원정보변경
연락처	017-234-	조성호
배송지	서울특별시 강남구	017-2**-
		20*****

결제금액	
결제금액	21,387원

그림 9-32 물건 구매 화면

■ 네이버 물건 구매

- 주문ID, 고객ID(조성호), 상품ID(시디즈), 수량(1), 구매 날짜(2017-10-25), 결제 방식(카드)으로 구성
- 상품 가격은 상품 데이터 베이스의 가격 필드에서 가져옴
- 배송지 정보는 고객 데이터의 주소 필드에서 받아옴
- 구매 스키마는 [주문ID, 고객ID, 상품ID, 수량, 구매날짜, 결제방식]으로 구성

주문ID	고객ID	상품ID	수량	구매날짜	결제 방식
201710	21248	7062345	1	2017 10 25	카드 결제

고객ID	아이디	비밀번호	이름	생년월일	성별	이메일	휴대전화
21248	SHCho	*****	조성호	1999 09 09	남	zoch@hs.ac.kr	017-234-21

그림 9-33 상품 스키마와 레코드

Thank you!