

통계데이터베이스 - 개요

Jinseog Kim
Dongguk University
jinseog.kim@gmail.com

2018-09-16

1 학습목표

- 통계소프트웨어 R을 이용한 다양한 데이터 조작을 할 수 있다.
- MySQL 쿼리문을 이용하여 데이터베이스 자료를 핸들링할 수 있다.
- 통계소프트웨어 R에서 Excel 및 MySQL 데이터베이스 자료를 분석할 수 있다.
- 빅데이터의 개념을 이해하고 R과 데이터베이스를 이용하여 빅데이터를 분석할 수 있다.

2 데이터베이스 개요

- 데이터베이스는 여러 응용 시스템들이 공유하여 이용할 수 있도록 통합하여 저장한 운영 데이터들의 집합
- 데이터베이스의 특징
 1. 최소한의 중복(minimal redundancy)만을 허용하는 통합 데이터(integrated data)
 2. 보조기억공간(디스크 등)에 저장하여, 필요 시에 검색하는 저장 데이터(stored data)
 3. 특정 단체의 작업 수행에 필요한 운영 데이터(operational data)
 4. 여러 사용자가 응용 프로그램, 응용 시스템을 통해 공동으로 사용하는 공용 데이터(shared data)
- 데이터베이스시스템(DBMS; Database Management System)은 자료들을 전산화된 레코드 형식으로 유지, 관리하여 주는 소프트웨어시스템
- DBMS의 활용 예
 1. 기업 : 제품 판매에 관한 데이터 유지
 2. 학교 : 학생에 관련된 (성적, 증명, 수업료 등) 정보 관리나 도서관의 목록, 대출 관리
 3. 항공사 : 비행 정보, 예약 상황 관리 등

2.1 데이터 모델

- 데이터 모델 (data model): 물리적 혹은 추상적인 현실 자료를 정형화된 형태로 표현하는 방식
- 관계형 데이터모델(relational data model): 테이블 형식을 이용하여 데이터를 정의하고 설명하는 데이터 모델
- 관계형데이터 베이스: 관계형 데이터 모델에 기반하여, 하나 이상의 테이블을 구현한 데이터베이스

2.2 데이터베이스 및 테이블 용어 정리

- 튜플(레코드) 유일성(uniquness) : 테이블 내에는 2개 이상의 동일 튜플이 존재할 수 없음
- 튜플들은 순서를 갖지 않음
- 속성(애트리뷰트)들은 순서를 갖지 않음
- 속성은 반드시 원자값(atomic value)를 가져야 함
- 키(key) : 테이블에서 특정 튜플을 식별 가능하게 하는 속성의 집합. 다음과 같은 종류가 있음
 - 슈퍼키(super key): 아무런 제약없이 튜플을 구분할 수 있는 속성의 집합, 예를 들면 주민등록번호, (학번, 주민등록번호)가 슈퍼키가 될 수 있음
 - 후보키(candidate key): 키 중 최소한의 속성만으로 구성된 키 (ex. 주민등록번호)
 - 기본키(primary key): 후보키중 하나만을 선택하여 식별자로 사용

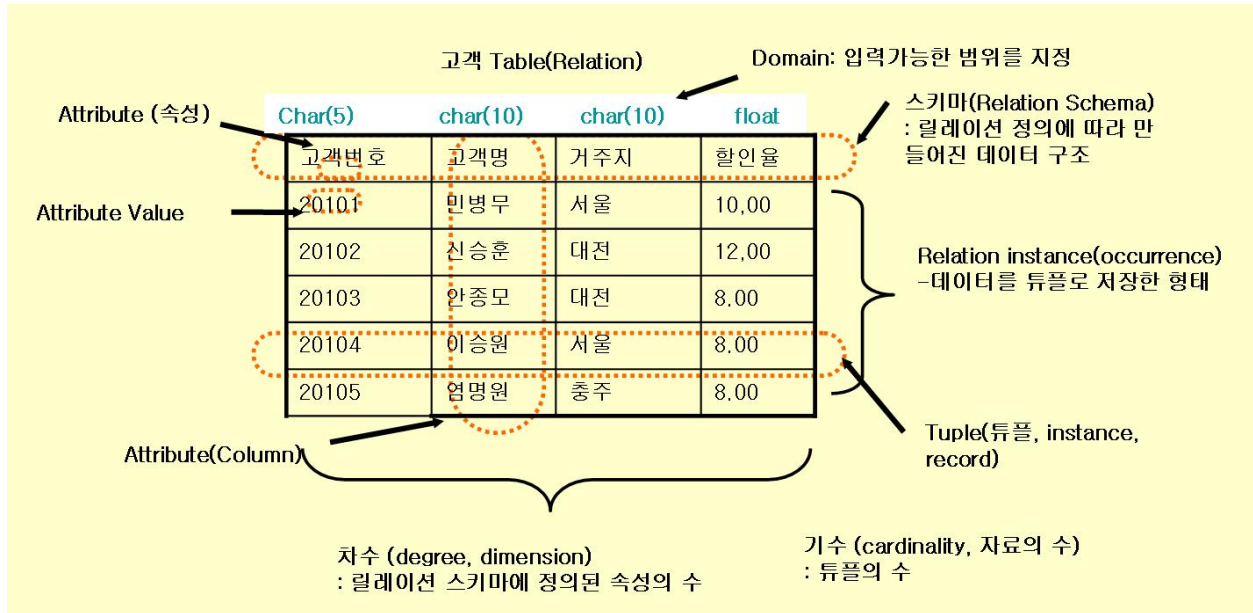


Figure 1: 관계형데이터모델에서의 용어

- 외래키(foreign key): 두 개 이상의 테이블에서 정의
- 테이블 스키마: 데이터베이스 관리 시스템에서 데이터 구조를 기술
- 데이터베이스 스키마: 데이터베이스에서 테이블 스키마의 집합

3 SQL 개요

- SQL(Structured Query Language) 이란 DBMS를 다루기 위한 표준언어
- 응용 프로그램과 DBMS 사이의 중간 다리 역할 담당
- SQL의 기능
 - DBMS를 제어
 - 데이터를 정의
 - 데이터를 조작

3.1 SQL 구성

- 데이터정의 언어(Data definition Language: DDL) : 데이터의 저장구조를 정의하는 언어로 스키마 정의 수정, 삭제하는 명령문과 인덱스의 생성 및 삭제
- 데이터조작 언어(Data Manipulation Language: DML) : 데이터에 접근 및 조작하는 언어로 튜플의 검색, 삽입, 삭제 수정과 같은 명령문

3.2 SQL 데이터베이스 제어문: 데이터베이스 객체 생성, 삭제, 갱신

- CREATE
- DROP
- ALTER
- CREATE DATABASE '데이터베이스_이름'

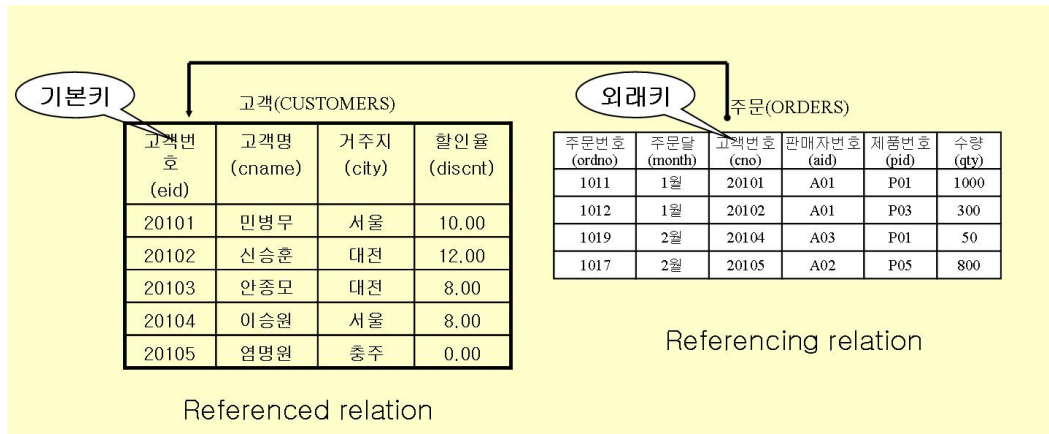


Figure 2: 외래키

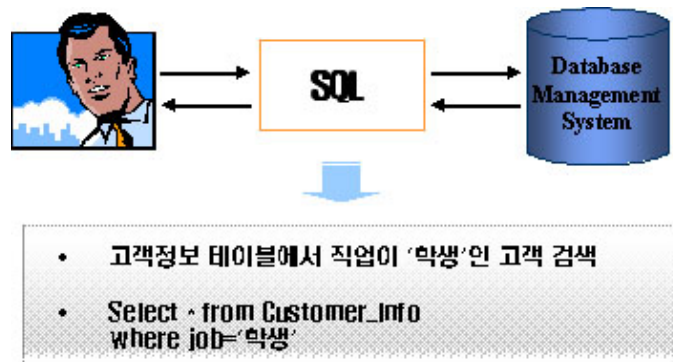


Figure 3: SQL의 역할

3.3 SQL 데이터 제어문: 데이터베이스 객체 이용권한 부여

- GRANT
- DENY
- REVOKE
- GRANT CREATE TABLE TO '사용자 계정'

3.4 SQL 데이터 조작문: 데이터 검색, 삽입, 갱신, 삭제

- SELECT
- INSERT
- UPDATE
- DELETE
- SELECT * FROM '테이블_이름'