

데이터베이스 개론

데이터 & 정보

➤ 데이터(data)

- 현실 세계에서 단순히 관찰하거나 측정하여 수집한 사실이나 값.

➤ 정보(information)

- 의사 결정에 유용하게 활용할 수 있도록 데이터를 처리한 결과물.
-

정보 처리

➤ 정보 처리(information processing)

- 데이터에서 정보를 추출하는 과정 또는 방법.

인터넷 쇼핑몰 주문 내역

주문 번호	주문 일자	제품명	판매 금액
1	2019-01-10	냉장고	50만 원
2	2019-02-12	세탁기	30만 원
3	2019-03-03	세탁기	30만 원
4	2019-04-05	에어컨	70만 원
5	2019-05-15	에어컨	80만 원
6	2019-06-19	에어컨	70만 원
7	2019-07-07	에어컨	70만 원
8	2019-08-12	냉장고	40만 원
9	2019-10-11	청소기	10만 원
10	2019-12-27	전자레인지	15만 원

데이터

정보 처리

제품별 총 판매액

제품	총 판매액
에어컨	290만 원
냉장고	90만 원
세탁기	60만 원
전자레인지	15만 원
청소기	10만 원

정보

분기별 총 판매액



데이터의 분류

➤ 정형 데이터 (structured data)

- 구조화된 데이터, 즉 미리 정해진 구조에 따라 저장된 데이터.
- 예 : 엑셀의 스프레드시트, 관계 데이터베이스의 테이블.

➤ 반정형 데이터 (semi-structured data)

- 구조에 따라 저장된 데이터이지만 데이터 내용 안에 구조에 대한 설명이 함께 존재.
- 구조를 파악하는 파싱(parsing) 과정이 필요하고, 보통 파일 형태로 저장.
- 예 : 웹에서 데이터를 교환하기 위해 작성하는 HTML, XML, JSON 문서나 웹 로그, 센서 데이터 등.

➤ 비정형 데이터 (unstructured data)

- 정해진 구조가 없이 저장된 데이터.
 - 예 : 소셜 데이터의 텍스트, 영상, 이미지, 워드나 PDF 문서와 같은 멀티미디어 데이터.
-

데이터베이스

➤ 정보 시스템(information system)

- 조직 운영에 필요한 데이터를 수집하여 저장해두었다가 필요할 때 유용한 정보를 만들어 주는 수단.

➤ 데이터베이스(DataBase)

- 기업이 지속적으로 유지 관리해야 하는 조직화된 정보의 집합.
- 정보 시스템 안에서 데이터를 저장하고 있다가 필요할 때 제공하는 역할을 담당.
- 특정 조직의 여러 사용자가 **공유**하여 사용할 수 있도록 **통합**해서 **저장**한 **운영** 데이터의 집합.

➤ 데이터베이스 관리 시스템(DataBase Management System)

- 연관성 있는 데이터들의 집합을 효율적으로 응용하기 위해 구성된 소프트웨어들의 집합체.
 - 방대한 양의 데이터를 편리하게 저장하고 효율적으로 관리하고 검색할 수 있는 환경을 제공해주는 시스템 소프트웨어.
 - 데이터를 공유하여 정보의 체계적인 활용 가능.
 - 응용 프로그램과 데이터베이스의 중재자로서 모든 응용 프로그램들이 데이터베이스를 공용할 수 있게끔 관리해 주는 소프트웨어 시스템.
-

데이터베이스 정의

➤ 공유 데이터(shared data)

- 여러 사용자들이 서로 다른 목적으로 사용하는 공유 가능한 데이터.

➤ 통합 데이터(integrated data)

- 데이터베이스는 똑같은 데이터가 원칙적으로 중복되어 있지 않다는 것.
- 데이터의 중복은 일반적으로 관리상의 복잡한 부작용을 초래.

➤ 저장 데이터(stored data)

- 컴퓨터가 접근할 수 있는 기억장치에 저장된 데이터.
- 주로 하드디스크에 저장되어 관리됨.

➤ 운영 데이터(operational data)

- 존재 목적이 명확하고 유용성을 지니고 있는 데이터.
 - 즉, 단순히 데이터를 모아둔 개념이 아닌 병원 관리를 위한 데이터 구축과 같은 목적이 분명한 데이터.
-

데이터베이스 특징

➤ 실시간 접근성(Real-time accessibility)

- 다수의 사용자의 요구에 대해서 처리 시간이 몇 초를 넘기지 말아야 한다는 의미.

➤ 지속적인 변화(Continuos evolution)

- 데이터베이스에 저장된 데이터는 지속적인 삽입, 삭제, 수정을 통해 최신의 정보가 정확하게 저장되어 처리되어야 함.

➤ 동시 공유(Concurrent sharing)

- 동일 데이터를 동시에 서로 다른 목적으로 사용할 수 있어야 함.

➤ 내용에 대한 참조(content reference)

- 데이터베이스 내에 있는 데이터 레코드들은 주소나 위치에 의해 참조되는 것이 아니라, 가지고 있는 값(내용)에 따라 참조해야 함.
 - 예) 재고량이 1,000개 이상인 제품의 이름 검색.
-

관계형 데이터베이스 (RDBMS: Relational DataBase Management System)

- 가장 일반적인 형태의 DBMS – 작성과 이용이 비교적 쉽고 확장이 용이.
- 오라클(Oracle), 사이베이스(Sybase), 인포믹스(Infomix), MYSQL, Acess, SQL Server.
- 관계형 데이터베이스 정보를 테이블 형태로 저장.
- 테이블은 2차원 형태의 표처럼 볼 수 있도록 로우(ROW:행)와 칼럼(COLUMN:열)으로 구성.

테이블 이름 : DEPT

DEPTNO	DNAME	LOC
10	ACCOUNTING	NEW YORK
20	RESEARCH	DALLAS
30	SALES	CHICAGO
40	OPERATIONS	BOSTON

- DEPARTMENT 테이블은 4개의 로우와 3개의 칼럼(부서번호:DNO, 부서이름:DNAME, 지역:LOC 으로 구성된 테이블.

관계형 데이터베이스 (RDBMS: Relational DataBase Management System)

➤ 데이터 딕셔너리(Data Dictionary : DD)

- 관계형 데이터베이스에서 객체를 정의하게 되면 그 객체가 가진 메타 데이터(metadata)의 정보가 저장되는 곳.
- 사용자에게 의해서 추가, 삭제, 수정되지 못하며 오로지 오라클 시스템에 의해서만 가능.

➤ SQL(Structured Query Language)

- 사용자와 관계형 데이터베이스를 연결시켜 주는 표준 검색 언어.
- 데이터베이스에 저장된 데이터를 조회, 입력, 수정 삭제하는 등의 조작이나, 테이블을 비롯한 다양한 객체(시퀀스, 인덱스 등)를 생성 및 제어하는 역할.

SQL(Structured Query Language)

➤ SQL의 종류

- 데이터 정의어(DDL)
 - 데이터베이스 관리자나 응용 프로그래머가 데이터베이스의 논리적 구조를 정의하기 위한 언어로서 데이터 사전(Data Dictionary)에 저장.
- 데이터 조작어(DML)
 - 데이터베이스에 저장된 데이터를 조작하기 위해 사용하는 언어로서 데이터 검색(Retrieval), 추가(Insert), 삭제>Delete), 갱신(Update) 작업 수행.
- 데이터 제어어(DCL)
 - 데이터에 대한 접근 권한 부여 등의 데이터베이스 시스템의 트랜잭션을 관리하기 위한 목적으로 사용되는 언어.

SQL과 데이터 조회하기

유 형	명 령 문
DQL : Data Query Language(데이터 질의어)	SELECT(데이터 검색시 사용)
DML : Data Manipulation Language(데이터 조작어) - 데이터 변경 시 사용	INSERT(데이터 입력) UPDATE(데이터 수정) DELETE(데이터 삭제)
DDL : Data Definition Language(데이터 정의어) - 객체 생성 및 변경 시 사용	CREATE(데이터베이스 생성) ALTER(데이터베이스 변경) DROP(데이터베이스 삭제) RENAME(데이터베이스 객체이름 변경) TRUNCATE(데이터베이스 저장 공간 삭제)
TCL : Transaction Control Language(트랜잭션 처리어)	COMMIT(트랜잭션의 정상적인 종료처리) ROLLBACK(트랜잭션 취소) SAVEPOINT(트랜잭션내에 임시 저장점 설정)
DCL : Data Control Language(데이터 제어어)	GRANT(데이터베이스에 대한 일련의 권한 부여) REVOKE(데이터베이스에 대한 일련의 권한 취소)

오라클 접속하기

- 데이터베이스 사용자는 오라클 계정(Account)이라는 용어와 같은 의미로 사용.
- 오라클을 설치하면 한 개 이상의 데이터베이스 권한을 갖는 디폴트(기본적인) 사용자가 존재.
- 오라클에서 제공되는 사용자 계정.

사용자계정	설 명
SYS	오라클 Super 사용자 계정이며 데이터베이스에서 발생하는 모든 문제들을 처리할 수 있는 권한을 가지고 있다.
SYSTEM	오라클 데이터베이스를 유지보수 관리할 때 사용하는 사용자 계정이며, SYS 사용자와 차이점은 데이터베이스를 생성할 수 있는 권한이 없으면 불완전 복구를 할 수 없다.
SCOTT	처음 오라클을 사용하는 사용자의 실습을 위해 만들어 놓은 연습용 계정이다.
HR	이 역시 오라클에 접근할 수 있도록 샘플로 만들어 놓은 사용자 계정이다.

오라클 접속하기

➤ Command 환경에서 SQL*Plus 로그인

- 데이터베이스 접속을 시도하면 오라클 데이터베이스를 사용할 수 있는 사용자인지를 검증하기 위해서 사용자 계정과 암호를 묻게 됩니다.

형식	SQLPLUS 사용자계정/암호
----	------------------

- 오라클을 설치하게 되면 기본적으로 생성되는 사용자 계정 중에서 hr을 사용합니다. hr의 패스워드는 tiger입니다.

예	sqlplus hr/tiger
---	------------------