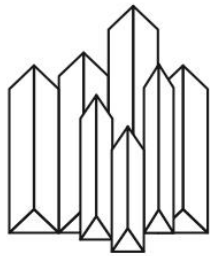


Chap01. 데이터베이스

- ▶ 01-1 데이터와 데이터베이스, DBMS
- ▶ 01-2 데이터 모델
- ▶ 01-3 관계형 데이터베이스와 SQL

01-1 데이터와 데이터베이스, DBMS

▶ 데이터와 정보



데이터(data)



정보(information)

데이터

A카드사에서 발급한 카드를 사용한
커피 전문점 결제 내역

정보

커피 전문점 결제 분포의 최상위
순위를 30~40대 남성이 차지했다는 결과

A카드사는 최근 몇 년간 급증한 커피 소비 동향을 파악하기 위해 A카드사에서 발급한 카드를 사용한 커피 전문점 결제 내역을 성별과 나이 대별로 분류하였다.

이 분류 작업과 관련하여 A카드사는 커피 전문점 결제 분포에서 20대 또는 30대 여성이 압도적으로 우위에 있을 것이라 예상했다.

그러나 결과는 뜻밖이었다. 커피 전문점 결제 분포의 최상위 순위를 30~40대 남성이 차지하고 있었던 것이다. 20~30대 여성의 결제 비율을 가볍게 넘어설 정도의 차이가 벌어진 것은 아니지만 예상을 뒤집는 결과였다.

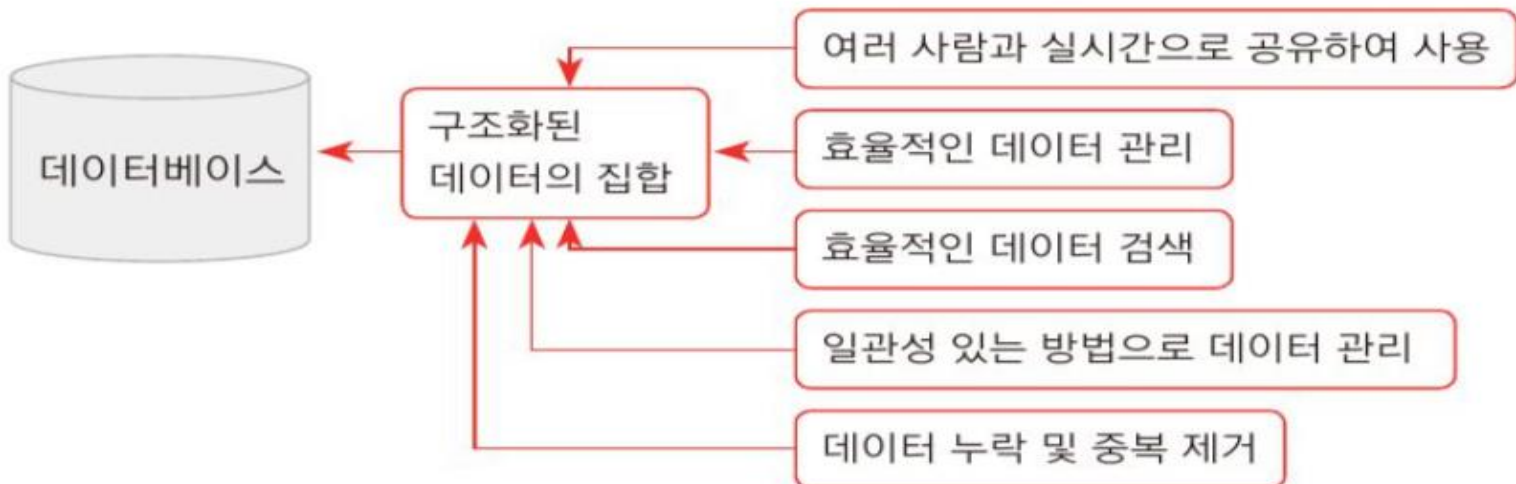
01-1 데이터와 데이터베이스, DBMS

▶ 데이터를 저장하는 여러가지 방법



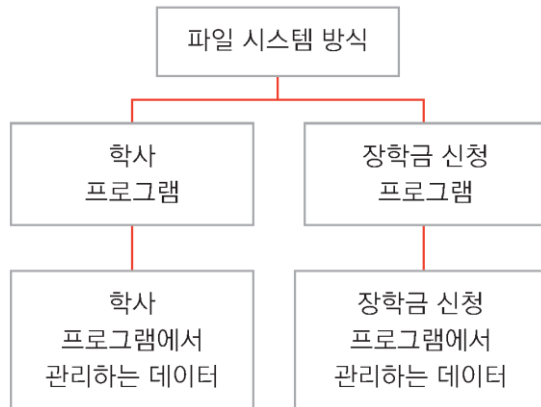
위 그림과 같이 여러가지 방법으로 수집한 데이터는 분석을 위한 통합 작업만으로도 시간과 비용이 많이 듭니다. 데이터가 여기저기 흩어져 있다면 최신 데이터를 정확하게 찾아내는 게 쉽지 않겠죠.

만일 데이터가 누락되거나 중복된다면 정확한 분석을 기대할 수 없고 결국 비싼 비용과 많은 시간을 투자한 분석이 실패하게 됩니다. 따라서 가치 있는 정보를 얻으려면 다음 조건에 맞게 데이터를 효율적으로 수집 · 통합하고 체계적으로 관리 · 분석해야 합니다.

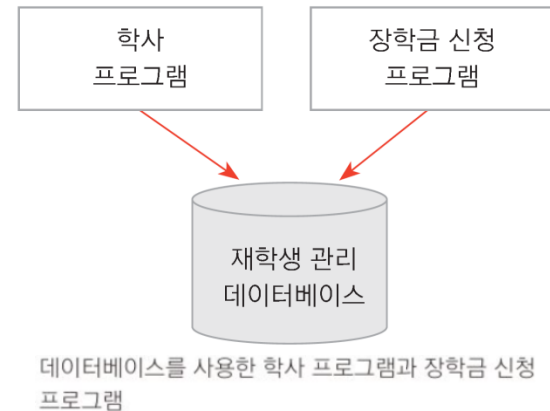


01-1 데이터와 데이터베이스, DBMS

▶ 파일 시스템 VS DBMS



파일 시스템의 예



학사 프로그램

학번	이름	학과	학년	학기	상태
2018-00001	홍길동	철학과	1	2	군휴학
2018-00002	이수선	컴퓨터공학과	4	2	졸업
2018-00003	이지수	경영학과	2	1	재학
2018-00004	김연아	사회체육학과	3	1	휴학

장학금 신청 프로그램

신청 장학금 종류	신청 일자	학번	재학 상태	장학금 신청 가능 여부
국가	20180409	2018-00001	군 휴학	신청 불가
성적	20180310	2018-00002	재학	신청 가능
동문	20180223	2018-00003	재학	신청 가능
근로	20180213	2018-00004	휴학	신청 불가

장학금 신청 프로그램에 졸업 정보가 누락된 경우

01-1 데이터와 데이터베이스, DBMS

▶ DBMS를 통한 데이터 관리



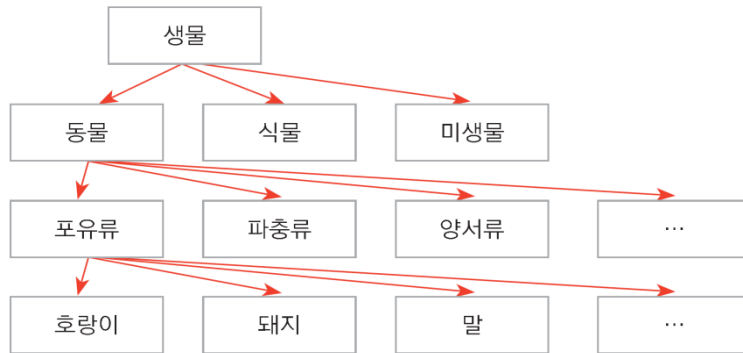
▶ 데이터의 접근 · 관리 등의 업무를 DBMS가 전담

파일 시스템 방식의 문제		DBMS를 통한 데이터 관리
데이터 중복	➔	하나의 소프트웨어가 데이터를 관리하므로 데이터 중복을 피할 수 있음
응용 프로그램이 개별 데이터를 직접 관리		여러 응용 프로그램이 하나의 DBMS를 통해 데이터를 사용하므로 데이터를 동시에 공유할 수 있음
응용 프로그램이 데이터를 쓰는 방식이 각각 다름		하나의 DBMS를 통해 데이터를 관리하기 때문에 각각의 응용 프로그램이 데이터를 관리하는 방식이 통합됨
데이터가 특정 응용 프로그램에 종속되어 있으므로 응용 프로그램을 변경하면 기존 데이터를 사용할 수 없음		응용 프로그램과는 별도로 데이터가 DBMS에 의해 관리 · 보관되기 때문에 응용 프로그램의 업데이트 또는 변경과 관계없이 데이터를 사용할 수 있음

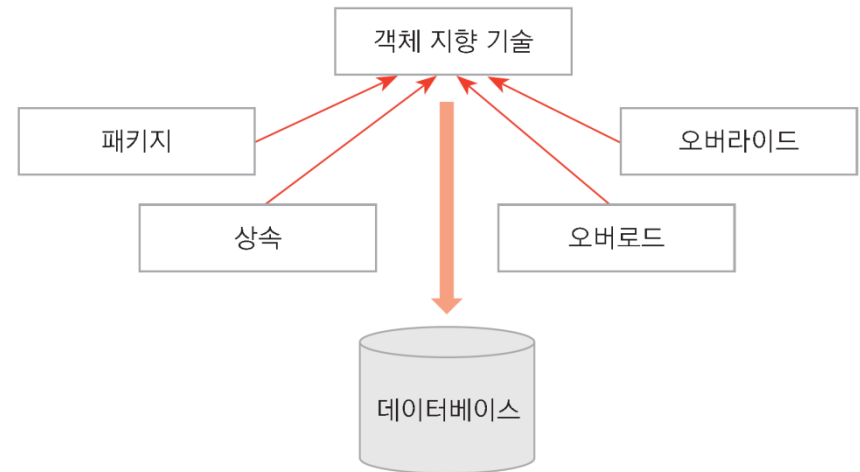
01-2 데이터 모델

▶ 계층형 vs 네트워크형 vs 객체지향형 데이터 모델

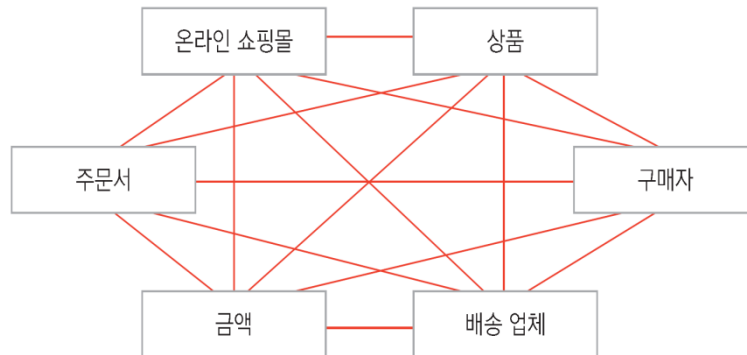
계층형 데이터 모델



객체지향형 데이터 모델



네트워크형 데이터 모델



01-2 데이터 모델

▶ 관계형 데이터 모델

사원 정보	사원 번호	사원 이름	사원 직급	부서 이름	위치
사원 번호	0001	홍길동	과장	회계	서울
사원 이름	0002	성춘향	대리	연구	인천
사원 직급	0003	박문수	사원	운영	분당
부서 이름	0004	심청이	사원	회계	서울
위치					

데이터 중복 발생

사원 정보와 부서 정보를 하나의 데이터로 관리할 경우

사원 정보	사원 번호	사원 이름	사원 직급	부서 코드
사원 번호	0001	홍길동	과장	10
사원 이름	0002	성춘향	대리	20
사원 직급	0003	박문수	사원	30
부서 코드	0004	심청이	사원	10

부서 정보	부서 코드	부서 이름	위치
부서 코드	10	회계	서울
부서 이름	20	연구	인천
위치	30	운영	분당

01-2 데이터 모델

▶ 관계형 데이터 모델

▶ 개체(entity), 속성(attribute), 관계(relationship)

이름	설명
개체 (entity)	데이터베이스에서 데이터화하려는 사물, 개념의 정보 단위입니다. 관계형 데이터베이스의 테이블(table) 개념과 대응되며 테이블은 릴레이션(relation)으로 표기하기도 합니다.
속성 (attribute)	개체를 구성하는 데이터의 가장 작은 논리적 단위로서 데이터의 종류·특성·상태 등을 정의합니다. 관계형 데이터베이스의 열(column) 개념과 대응됩니다.
관계 (relationship)	개체와 개체 또는 속성 간의 연관성을 나타내기 위해 사용합니다. 관계형 데이터베이스에서는 테이블 간의 관계를 외래키(foreign key) 등으로 구현하여 사용합니다.

객체(entity)

속성(attribute)

관계(relationship)

사원 정보	사원 번호	사원 이름	사원 직급	부서 코드
사원 번호	0001	홍길동	과장	10
사원 이름	0002	성춘향	대리	20
사원 직급	0003	박문수	사원	30
부서 코드	0004	심청이	사원	10

부서 정보	부서 코드	부서 이름	위치
부서 코드	10	회계	서울
부서 이름	20	연구	인천
위치	30	운영	분당

01-3 관계형 데이터베이스와 SQL

- ▶ 관계형 데이터베이스란?
 - ▶ 관계형 데이터 모델
 - ▶ RDBMS(Relational Database Management System)
 - ▶ Oracle, MS-SQL, MySQL, MariaDB, PostgreSQL, DB2, etc...

343 systems in ranking, June 2018								
Rank			DBMS	Database Model	Score			
Jun 2018	May 2018	Jun 2017			Jun 2018	May 2018	Jun 2017	
1.	1.	1.	Oracle +	Relational DBMS	1311.25	+20.84	-40.51	
2.	2.	2.	MySQL +	Relational DBMS	1233.69	+10.35	-111.62	
3.	3.	3.	Microsoft SQL Server +	Relational DBMS	1087.73	+1.89	-111.23	
4.	4.	4.	PostgreSQL +	Relational DBMS	410.67	+9.77	+42.13	
5.	5.	5.	MongoDB +	Document store	343.79	+1.67	+8.79	
6.	6.	6.	DB2 +	Relational DBMS	185.64	+0.03	-1.86	
7.	7.	↑ 9.	Redis +	Key-value store	136.30	+0.95	+17.42	
8.	↑ 9.	↑ 11.	Elasticsearch +	Search engine	131.04	+0.60	+19.48	
9.	↓ 8.	↓ 7.	Microsoft Access	Relational DBMS	130.99	-2.12	+4.44	
10.	10.	↓ 8.	Cassandra +	Wide column store	119.21	+1.38	-4.91	

RDBMS는 세계 DBMS 시장에서 매우 높은 점유율을 차지하고 있습니다. 그중에서도 오라클 데이터베이스 순위는 1위입니다.

01-3 관계형 데이터베이스와 SQL

▶ SQL이란?

▶ Structured Query Language

종류	설명
DQL(Data Query Language)	RDBMS에 저장한 데이터를 원하는 방식으로 조회하는 명령어
DML(Data Manipulation Language)	RDBMS 내 테이블의 데이터를 저장·수정·삭제하는 명령어
DDL(Data Definition Language)	RDBMS 내 데이터 관리를 위해 테이블을 포함한 여러 객체를 생성·수정·삭제하는 명령어
TCL(Transaction Control Language)	트랜잭션 데이터의 영구 저장·취소 등과 관련된 명령어
DCL(Data Control Language)	데이터 사용 권한과 관련된 명령어