

Chapter 1: 데이터베이스 환경

Institute of Management Research, PNU

Dr. Jinsung Kim

kimsam@pusan.ac.kr

“매니저 나와”

- 고등학생 딸에게 유아용품 할인쿠폰 발송
- 미국 대형마트 ‘타겟(Target)’의 미니애폴리스(Minneapolis) 지점 사례
- 대형마트 마케팅 팀은 부모도 몰랐던 딸의 임신 사실을 어떻게 알아챘을까?

들어가기

- 온라인 쇼핑 결제 건수가 가장 많은 요일, 지역은?
월, 화요일, 여성 1인 가구가 많은 강남구
- 백화점과 대형마트에 가장 손님이 많은 요일은?
금요일 → 목요일 (법인카드 결제 비중이 높은 요일)
- 우리(기업)는 알고 싶다.
 - 누가 무엇을 언제 얼마만큼 샀는지? 그것도 빠르고 정확하게...
- 당신은 고객 관리를 위해 어떤 프로그램을 사용하십니까?
워드? 엑셀?

들어가기

고객

고객ID	고객이름	주민번호	전화번호
1	백현	920506-1xxxxxx	111-1xxx
10	카이	940114-1xxxxxx	123-1xxx
11	레이	911007-1xxxxxx	234-2xxx
12	찬열	921127-1xxxxxx	345-3xxx
13	지민	910108-2xxxxxx	456-4xxx
14	초아	900306-2xxxxxx	567-5xxx
15	유아	921230-2xxxxxx	678-6xxx
2	루한	900420-1xxxxxx	222-2xxx
3	디오	930112-1xxxxxx	333-3xxx
4	타오	930502-1xxxxxx	444-4xxx
5	첸	920921-1xxxxxx	555-5xxx
6	세훈	940412-1xxxxxx	666-6xxx
7	수호	910522-1xxxxxx	777-7xxx
8	크리스	901106-1xxxxxx	888-8xxx
9	시우민	900326-1xxxxxx	999-9xxx

제품

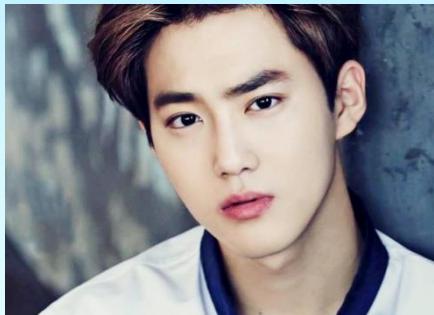
제품ID	제품명	표준가격
1	비디오	175
2	오디오	200
3	라디오	375
4	노트북	650
5	PC	725
6	프린터	350
7	스캐너	150
8	프로젝트	450

- 고객이 주문내역을 어떻게 관리(저장)할 것인가?

들어가기

주문내역

고객_ID	제품_ID	주문날짜	주문수량
1	2	21-Oct-08	1
1	4	21-Oct-08	1
1	6	21-Oct-08	1
7	1	22-Oct-08	1
7	2	22-Oct-08	2
7	3	22-Oct-08	1
14	5	25-Oct-08	1
14	7	25-Oct-08	1



BOSS가 당신에게 묻습니다.

- 초아(수호)는 언제 무엇을 얼마만큼 주문했나?
- 누가 가장 많이 주문했나?
- 언제 가장 많은 주문이 있었나?
- 어떤 제품이 가장 많은 주문을 받았나?

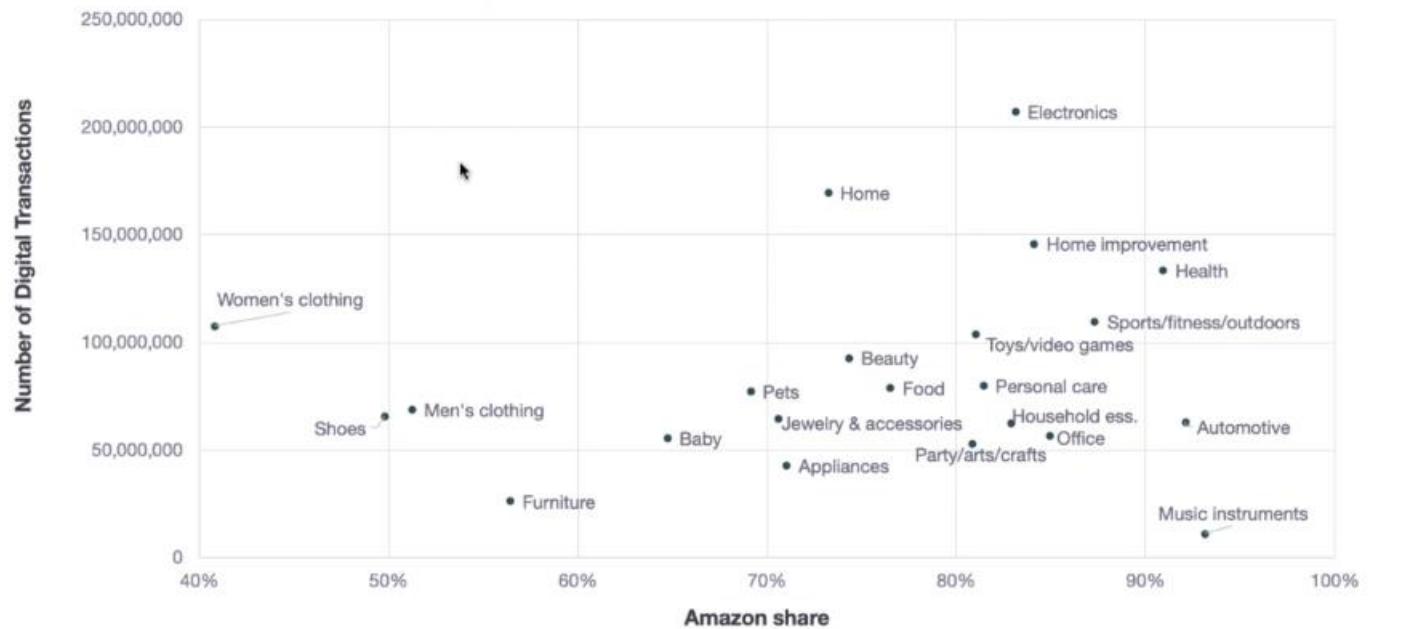
아마존 북스

- 2017.05.25 미국 뉴욕 맨하탄에 'amazon books' 오픈
 - 112평 규모로 약 3,000권의 책을 판매
 - 미국 최대 서점 '반스앤노블'이 사업 실패로 철수한 지역



아마존 매출 분야

Mapping the eCommerce Space: Category Size & Share



프라임 워드로브(Prime Wardrobe)서비스 실시: 옷을 직접 입어보고 반품 or 구매 결정

아마존의 AWS

〈표 1〉 2016–2017년 전 세계 IaaS 퍼블릭 클라우드 서비스 시장점유율

(단위: 백만 달러)

2017년 순위	업체명	2016년 매출	2017년 매출	2017년 시장점유율	2016–2017 성장률(%)
1	아마존	9,775	12,221	51.8%	25.0%
2	마이크로소프트	1,579	3,130	13.3%	98.2%
3	알리바바	670	1,090	4.6%	62.7%
4	구글	500	780	3.3%	56.0%
5	IBM	297	457	1.9%	539%
	기타	5,392	5,902	25.0%	9.5%
-	총계	18,213	23,580	100.0%	29.5%

출처: Gartner(2018.6)

- 글로벌 시장조사 업체 시너지 리서치 그룹(Synergy Research Group)의 조사에 따르면, 아마존은 전 세계의 클라우드 인프라 서비스 시장의 34%를 차지.
- 2015년 말에 작성된 위치리스크 문서에 의하면, 아마존은 전 세계 9개국 15개 도시에 걸쳐 설치된 100개 이상의 데이터 센터를 비밀 운영(CIA, 록히드마틴, 국방성과 계약).

정의

- 데이터베이스(database): 논리적으로 연관된 데이터들의 구조화된 집합
- 데이터(data): 객체 및 이벤트에 관련된 사실로서 컴퓨터 매체에 기록되고 저장될 수 있는 것
 - 구조적: 숫자, 문자, 날짜
 - 비구조적: 사진, 비디오, 문서
- 빅데이터(big data): extremely large data sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions.
- 정보(information): 데이터 사용자의 지식을 증대시킬 수 있는 방법으로 가공된 데이터
- 메타데이터(metadata): 최종 사용자가 이용하는 데이터의 속성 또는 특성과 사용 정황을 기술한 데이터

그림 1-1 데이터의 정보로의 변화: (a) 상황과 연계된 데이터

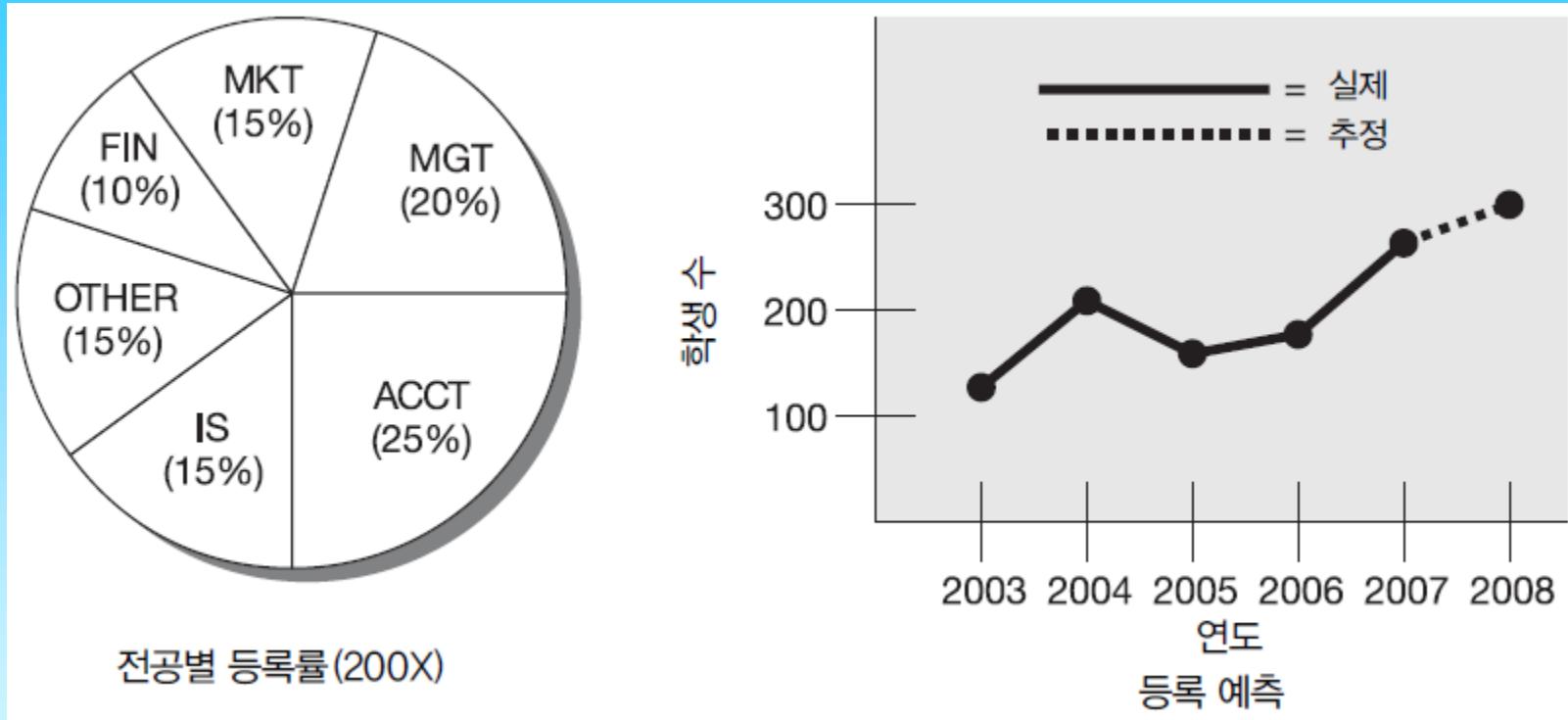
수업 명부

과목: MGT 500, 경영정책론
분반: 2

학기: 봄, 200X

이름	학번	전공	평점
Baker, Kenneth D.	324917628	MGT	2.9
Doyle, Joan E.	476193248	MKT	3.4
Finkle, Clive R.	548429344	PRM	2.8
Lewis, John C.	551742186	MGT	3.7
McFerran, Debra R.	409723145	IS	2.9
Sisneros, Michael	392416582	ACCT	3.3

그림 1-1 데이터의 정보로의 변화: (b) 요약된 데이터



그래픽은 데이터를 유용한 정보로 바꾸어
보여주며 이를 관리자는 데이터 해석과 중요한
의사결정에 사용할 수 있다.

표 1-1 수업명부에 대한 메타데이터 예

데이터 항목		메타데이터				
이름	유형	길이	최소치	최대치	설명	원천
과목	Alphanumeric	30			과목코드와 이름	교무과
분반	Integer	1	1	9	분반 번호	학적과
학기	Alphanumeric	10			학기와 연도	학적과
이름	Alphanumeric	30			학생 이름	학생 정보시스템
학번	Integer	9			학생 번호	학생 정보시스템
전공	Alphanumeric	4			학생 전공	학생 정보시스템
평점	Decimal	3	0.0	4.0	학생 학점의 평균	교무과

데이터 항목들에 대한 메타 데이터로서 데이터 항목의 이름, 데이터 유형, 길이, 허용 값의 최소치와 최대치, 간단한 설명, 데이터 원천을 설명한다.

입고전표에 대한 ERD

입고전표

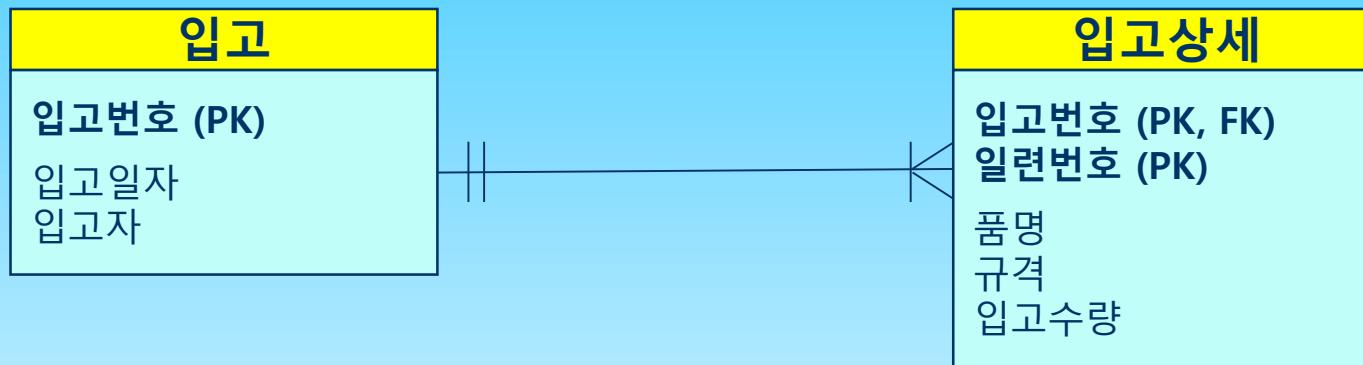
No. 234

입고일자 : 2019. 09. 15

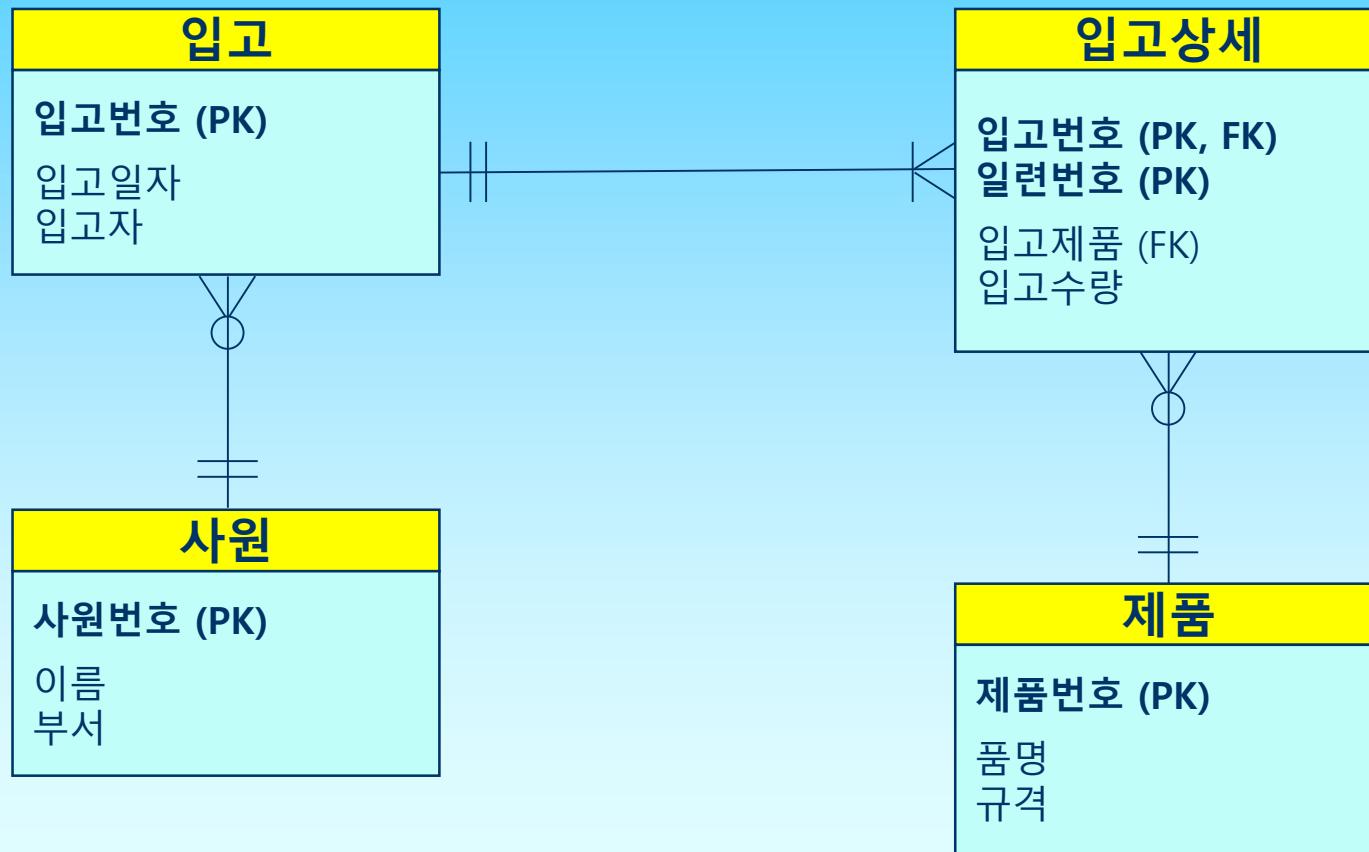
입고자 : 류현진

No	품명	규격	입고수량
1	냉장고	200L	20
2	세탁기	100x5	15

입고전표에 대한 ERD

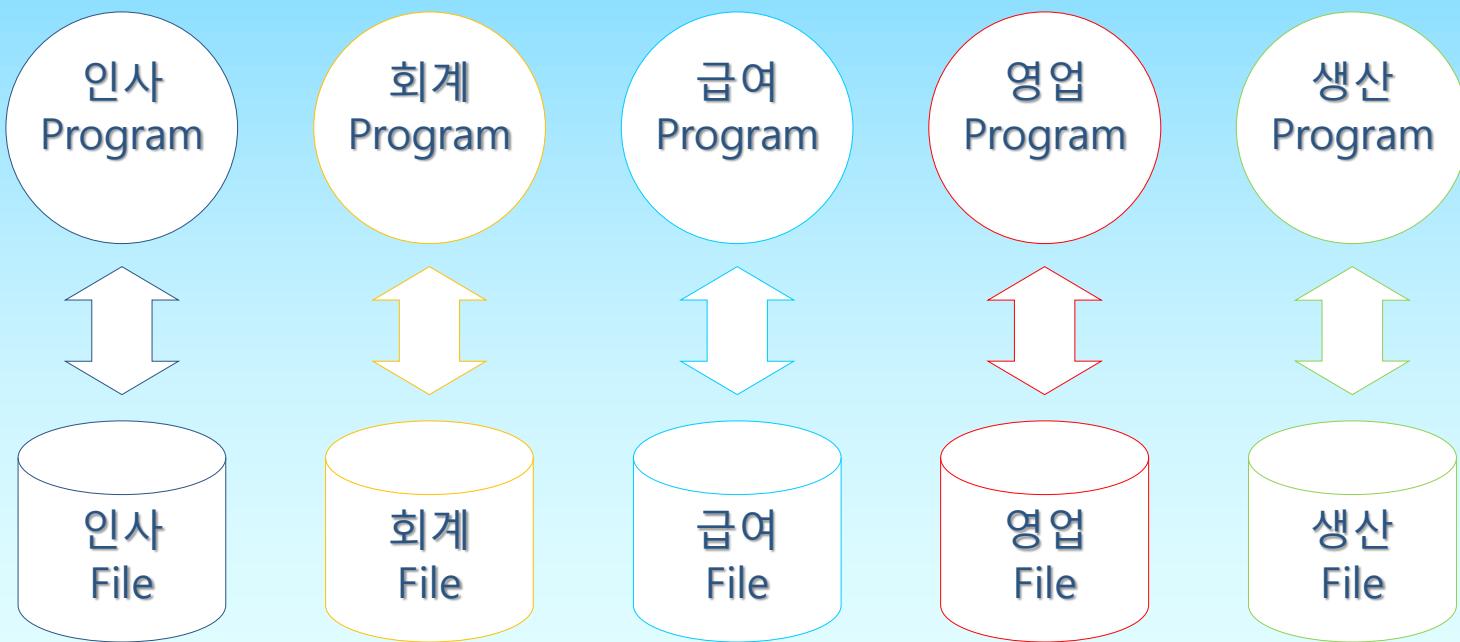


입고전표에 대한 ERD



파일처리시스템의 단점

- 파일 시스템은 각각의 파일 단위로 업무와 관련한 데이터를 저장하며 파일들을 처리하기 위한 독립적인 애플리케이션과 상호 연동되어 있다.



파일처리시스템의 단점

- **프로그램-데이터 종속**
 - 모든 프로그램들은 자신들이 사용하는 각 파일에 대한 메타 데이터를 유지한다.
- **데이터 중복**
 - 다른 시스템/프로그램들은 동일한 데이터의 별도의 복사본을 가지고 있다.
- **제한된 데이터 공유**
 - 데이터의 중앙집중식 제어가 없다.
- **긴 개발시간**
 - 프로그래머들은 자신만의 파일 형식을 설계해야 한다.
- **과도한 프로그램 유지보수**
 - 정보시스템 개발에 총 예산의 80% 정도 투입

데이터 종속의 문제

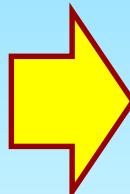
- 각 응용 프로그래머는 자신만의 데이터를 유지해야 한다.
- 각 응용 프로그램은 각 파일의 메타데이터 코드를 필요로 한다.
- 각 응용 프로그램은 데이터를 읽기, 생성, 업데이트, 삭제를 위한 자신만의 처리 루틴이 있어야 한다.
- 중앙 통제와 조정에 대한 부족
- 비표준 파일 형식

데이터 종속의 문제

- 데이터를 사용하는 프로그램 구조가 데이터 구조(파일 구조)의 영향을 받는다.

저장된 학생 정보 파일

98001손흥민	111-xxxx	27
98002류현진	123-xxxx	32
98003이강인	456-xxxx	18
98004추신수	337-xxxx	37
98005고진영	159-xxxx	24



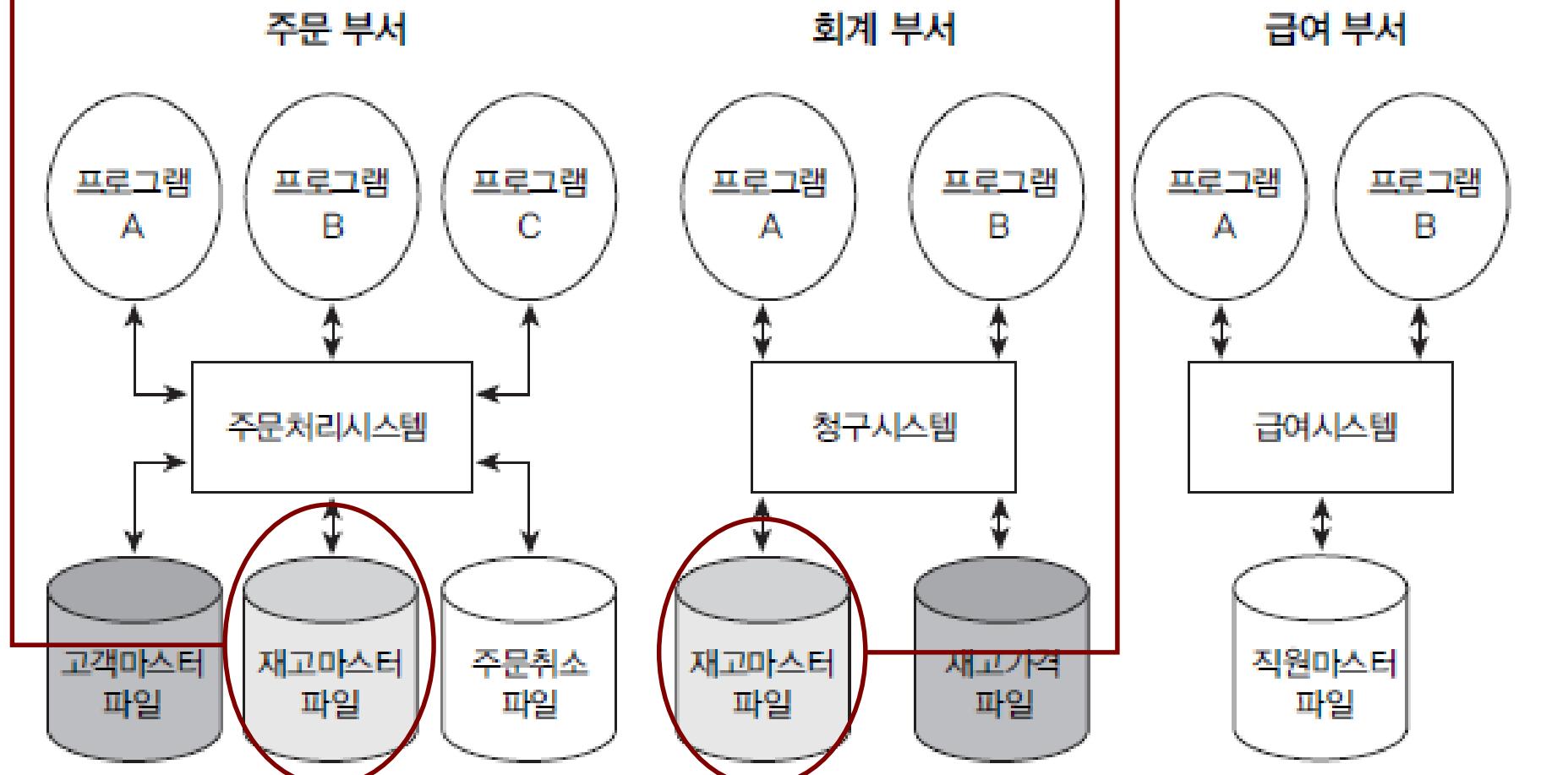
응용 프로그램

01 STUDENT-TBL	
03 STUDENT-RECORD	
05 S-ID	PIC X(05)
05 S-NAME	PIC X(20)
05 S-TEL	PIC X(10)
05 S-AGE	PIC X(02)

저장된 학생 정보를 읽어 처리하는 COBOL 프로그램
필드의 길이를 20자리에서 30자리로 늘려야 할 경우?

그림 1-3 Pine Valley Furniture Company의 오래된 파일처리 시스템

중복 데이터



데이터 중복의 문제

- 종복 데이터를 가지는 공간의 낭비
- 더 많은 유지 보수 문제의 원인
- 가장 큰 문제:
 - 하나의 파일에서의 데이터 변경은 불일치의 원인
 - 데이터 무결성 문제가 발생

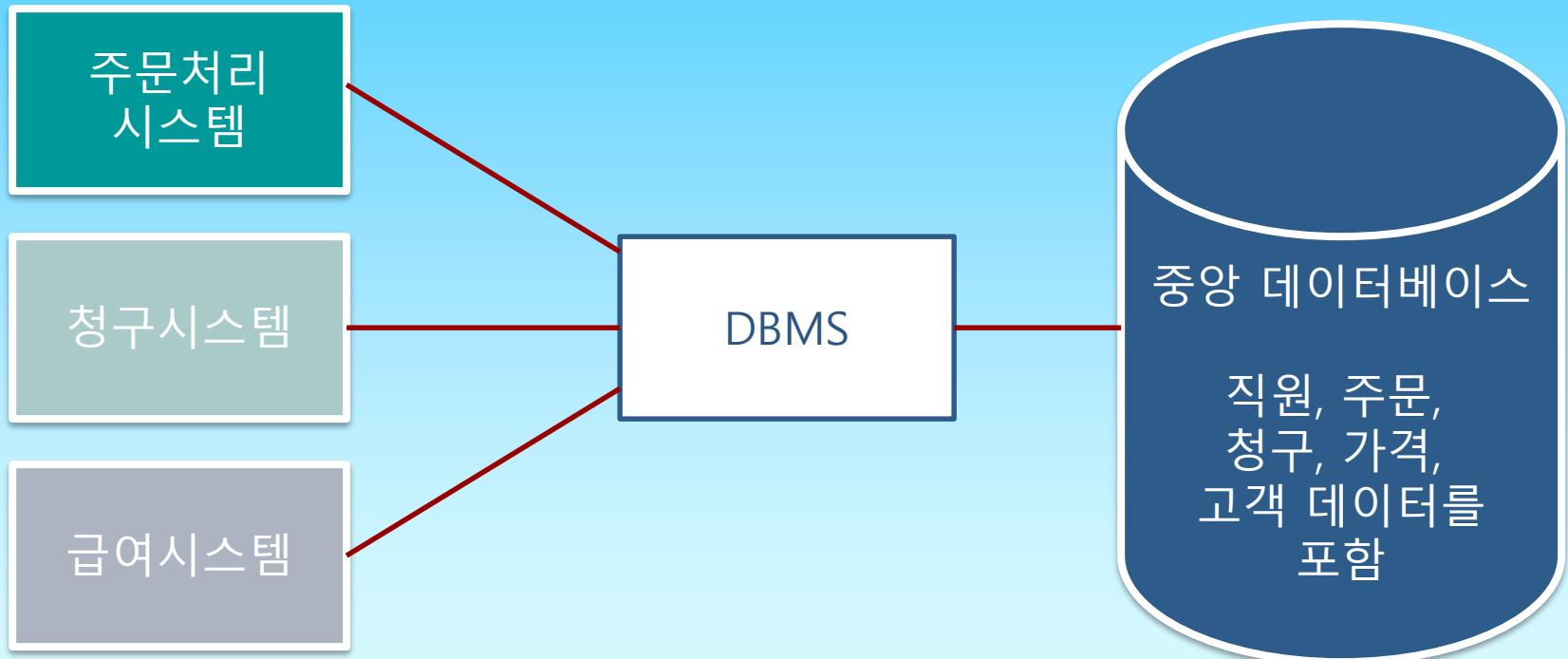
해결책: 데이터베이스 접근법

- 공유 데이터의 중앙 저장소
(repository)
- 데이터는 제어 에이전트에 의해 관리
- 표준화, 편리한 형태로 저장

Database Management System(DBMS) 필요

데이터베이스 관리 시스템

- 데이터베이스 생성, 유지 및 사용자 데이터베이스 접근을 제어하는데 사용되는 소프트웨어 시스템

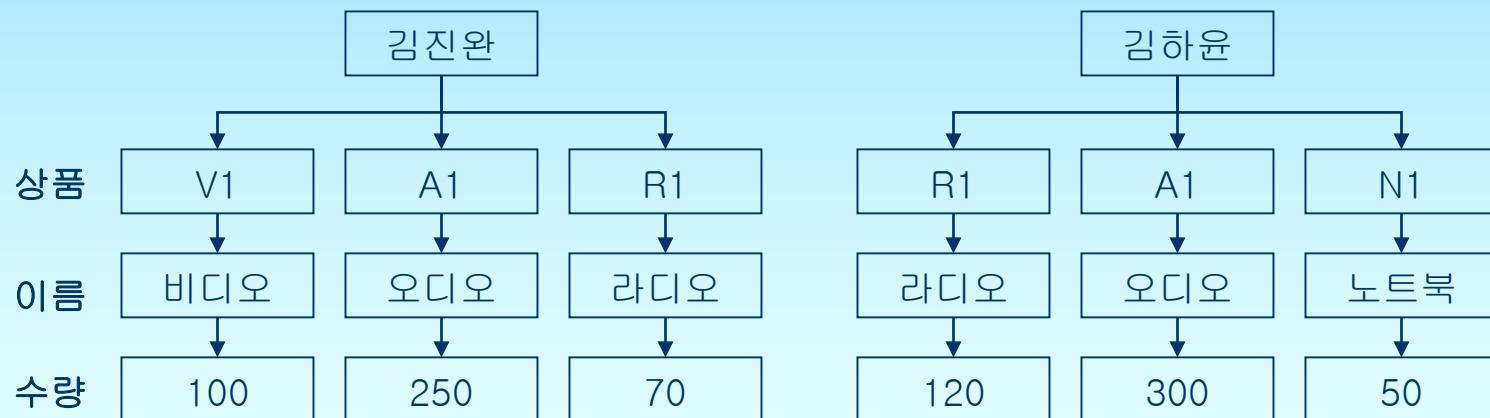


DBMS 는 하드웨어 자원을 관리하는 운영체제와 같이 데이터 자원을 관리한다.

데이터베이스 관리 시스템의 유형

■ 계층형(Hierarchical) DBMS

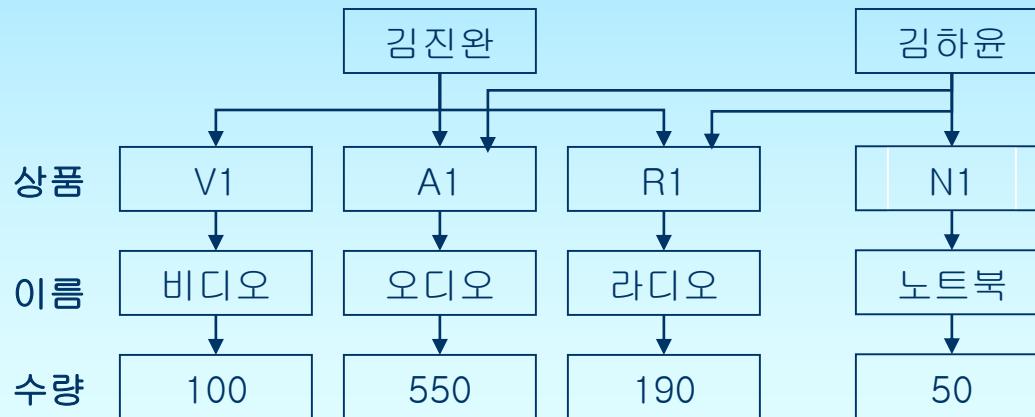
- 계층형 데이터베이스 관리시스템은 계층적으로 데이터가 상하 종속적인 관계로 구성
- 데이터의 액세스 속도가 빠르고 데이터 사용량을 쉽게 예측 가능
- 변화하는 프로세서에 대한 적응이 어려움



데이터베이스 관리 시스템의 유형

■ 네트워크형(Network) DBMS

- 데이터 구조를 네트워크상의 노드 형태로 논리적으로 표현한 데이터 모델로서 각각의 노드를 서로 대등한 관계로 구성한 시스템
- 구성과 설계가 복잡하고 데이터 종속성을 해결하지 못한 시스템



데이터베이스 관리 시스템의 유형

■ 관계형(Relational) DBMS

- 1967년 IBM의 연구원으로 있던 E. F. Codd가 수학적 기초에 근거하여 고안
- 데이터베이스는 최소한의 의미를 가지는 테이블들로 구성되며 각 테이블들은 연관성 있는 필드들로 연결
- 개체(Entity)와 릴레이션(Relation)을 중심으로 기업의 정보 구조와 업무 프로세스를 정의

고객번호	고객이름
1	김진완
10	김하윤
33	김예찬
45	김예서

상품번호	상품이름
V1	비디오
A1	오디오
R1	라디오
N1	노트북

고객이름	상품번호	수량
1	V1	100
10	A1	300
1	A1	250
45	N1	80
1	R1	70
10	R1	120
10	N1	50

데이터베이스 관리 시스템의 유형

- 객체지향형(Object-Oriented) DBMS
 - 멀티미디어 데이터의 원활한 처리를 위해 고안된 데이터베이스이며 객체 지향적 프로그래밍의 개념을 도입해서 구성한 시스템
- 객체관계형(Object-Relational) DBMS
 - 관계형 데이터베이스의 안정된 성능에 기반하면서 멀티미디어 데이터를 원활히 처리하기 위해 현실적으로 제안된 모델

데이터베이스 접근법의 장점들

- 프로그램-데이터 독립성
- 계획된 데이터 중복성
- 데이터 일치성 향상
- 데이터 공유 향상
- 업무 변화에 대한 적응 능력
- 애플리케이션 개발의 생산성 향상
- 표준 적용 강화
- 데이터 품질 향상
- 데이터 접근성 및 반응성 향상
- 프로그램 유지보수 노력의 감수
- 의사결정 지원 향상

데이터베이스 접근법의 비용 및 위험

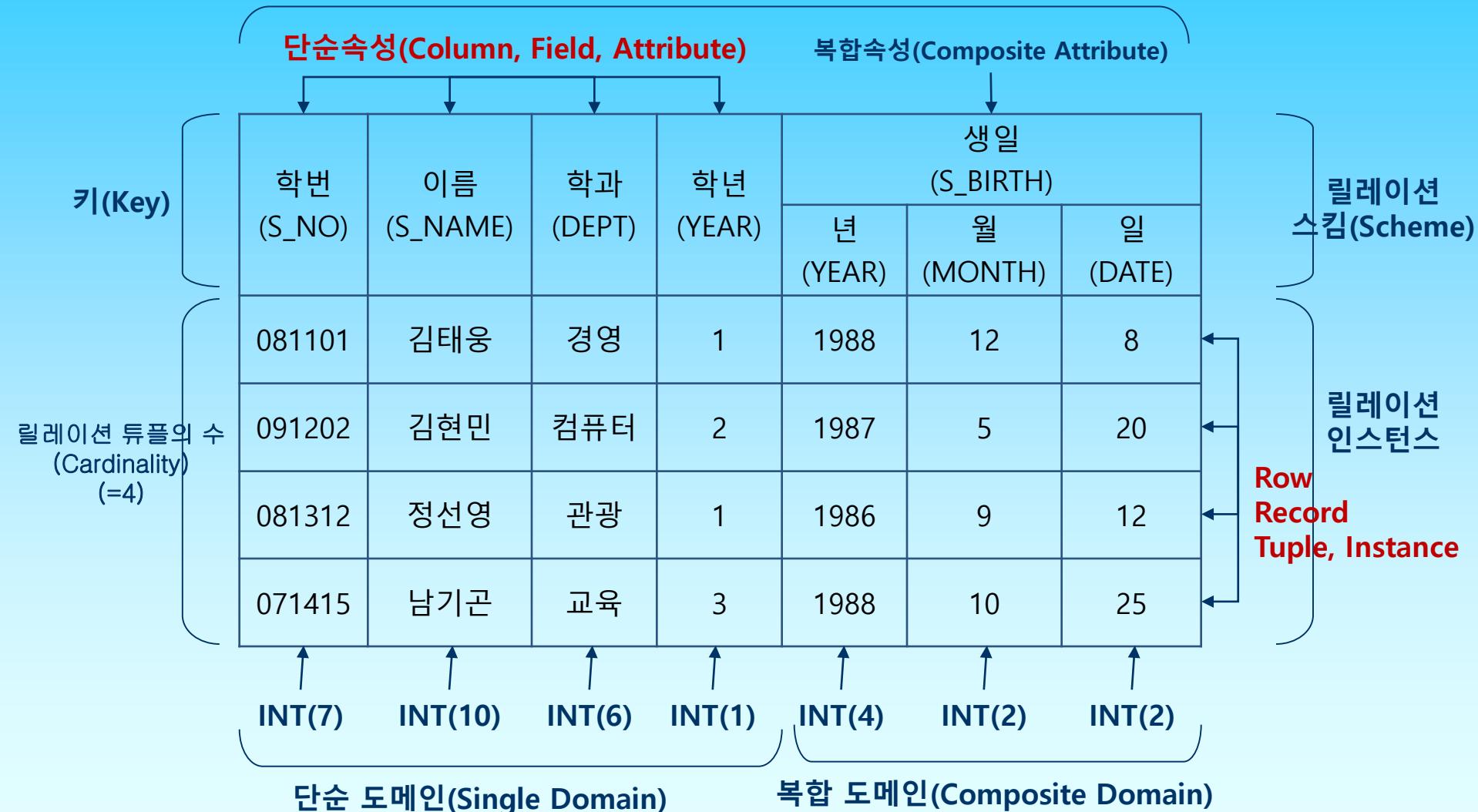
- 신규 전문 인력
- 설치 및 관리에 대한 비용 및 복잡성
- 전환 비용
- 확실한 백업 및 복구에 대한 필요성
- 조직 갈등

데이터베이스 접근법의 구성요소

- 데이터 모델 (Data models)
 - 데이터의 특질과 데이터들 간의 관계를 포착
 - 전사적 데이터 모델: 조직에 대한 상위 수준의 개체들과 그들 간의 관계성들을 보여주는 도식적 모델
 - 프로젝트 데이터 모델: 구체적인 내역들을 보여주며 데이터베이스의 데이터 구조와 일치하는 데이터 구조
- 관계형 데이터베이스 (Relational Databases)
 - 기본키/보조키의 관계를 보여주며 객체들을 나타내는 테이블(릴레이션)를 포함하는 데이터베이스 기술
- 인터넷 기술 사용
 - 네트워크와 정보통, 분산 데이터베이스, 클라이언트-서버, 3-계층 구조(클라이언트, 응용, 데이터 계층)
- 데이터베이스 어플리케이션
 - 어플리케이션 프로그램들은 데이터베이스 사용자들을 위하여 데이터베이스 행동(생성, 읽기, 업데이트, 삭제)을 수행하기 위해 사용

Relation / Entity

릴레이션 속성의 차수(Degree) (=7)



학생과 과정 관계를 표현한 ERD (Entity Relationship Diagram)

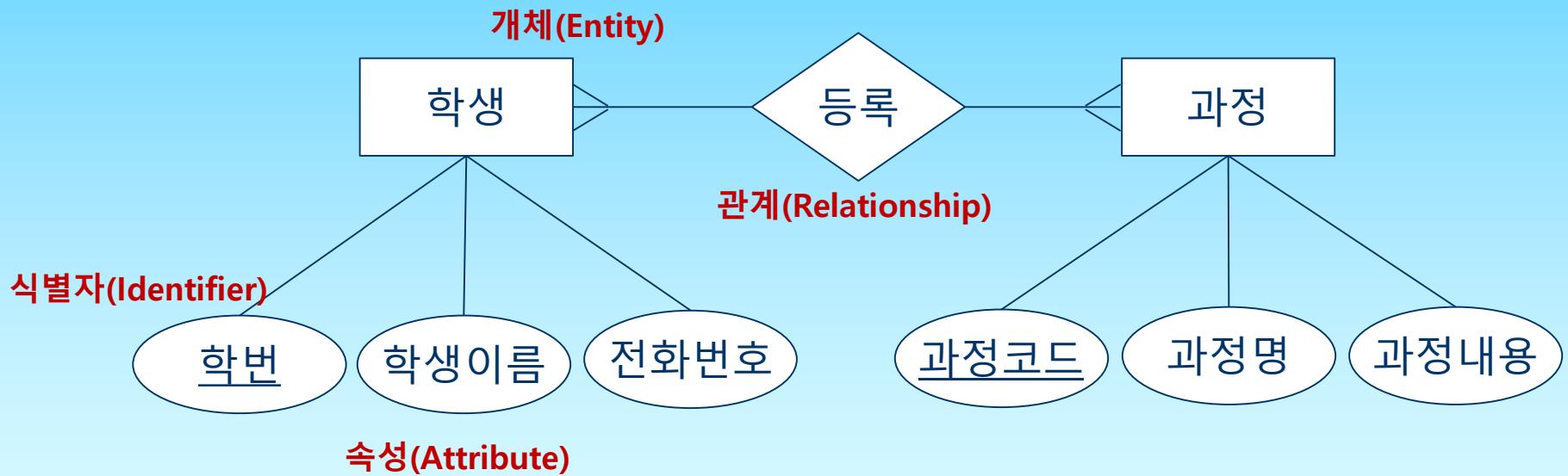
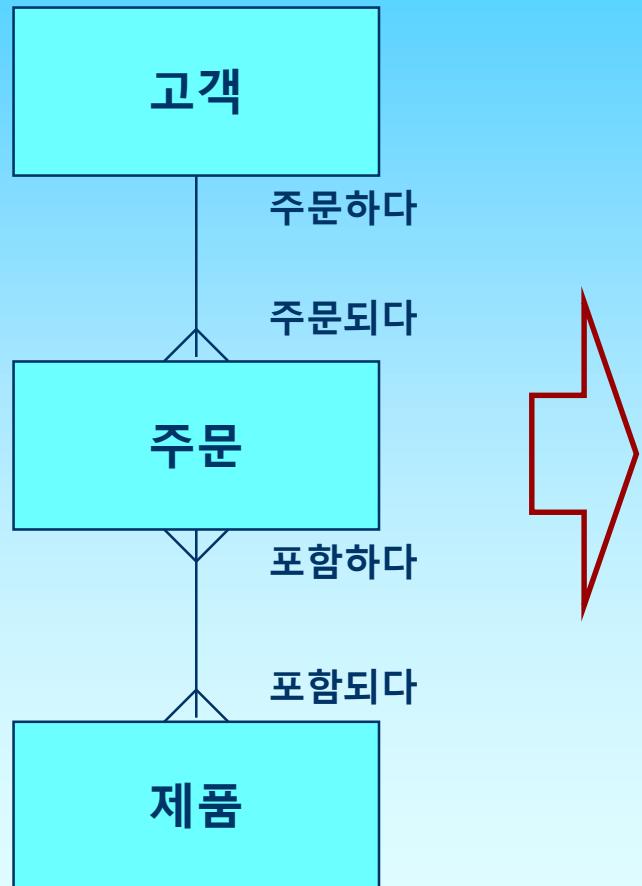
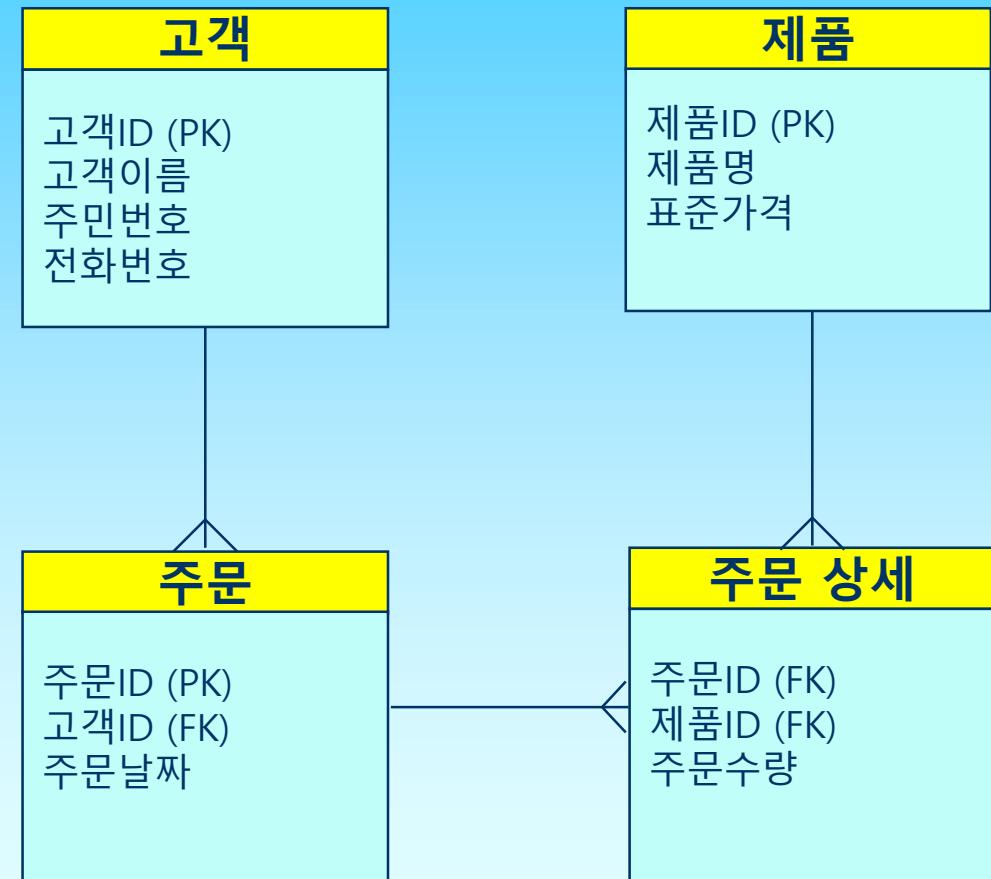


그림 1-2 전사적 수준의 데이터 모델과 프로젝트 수준의 데이터 모델 비교

전사적 데이터 모델의 일부



프로젝트 데이터 모델의 일부



고객

고객ID
고객이름
주민번호(AK)
전화번호

주문하다

주문

주문ID
고객ID(FK)
주문날짜

주문되다

한 고객은 여러
개를 주문할 수
있지만 각
주문은 한
고객에게 속한다.

→ 일대다 관계

제품

제품ID
제품명(AK)
표준가격

가지다

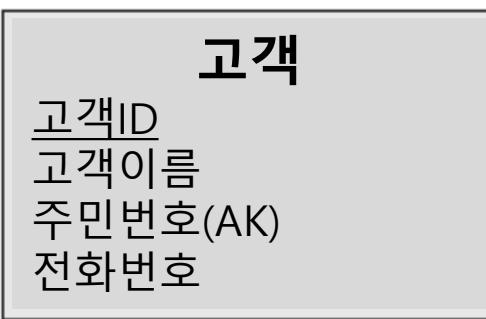
주문상세

주문ID(FK)
제품ID(FK)
주문수량

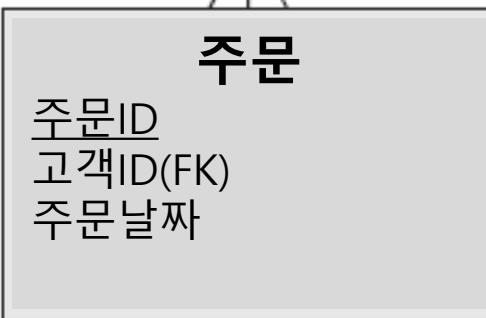
관련되다

포함되다

포함하다



주문하다



주문되다

한 주문은 여러 주문상세를 가지며; 각 주문상세는 하나의 주문과 관련된다.



가지다



관련되다

→ 일대다 관계

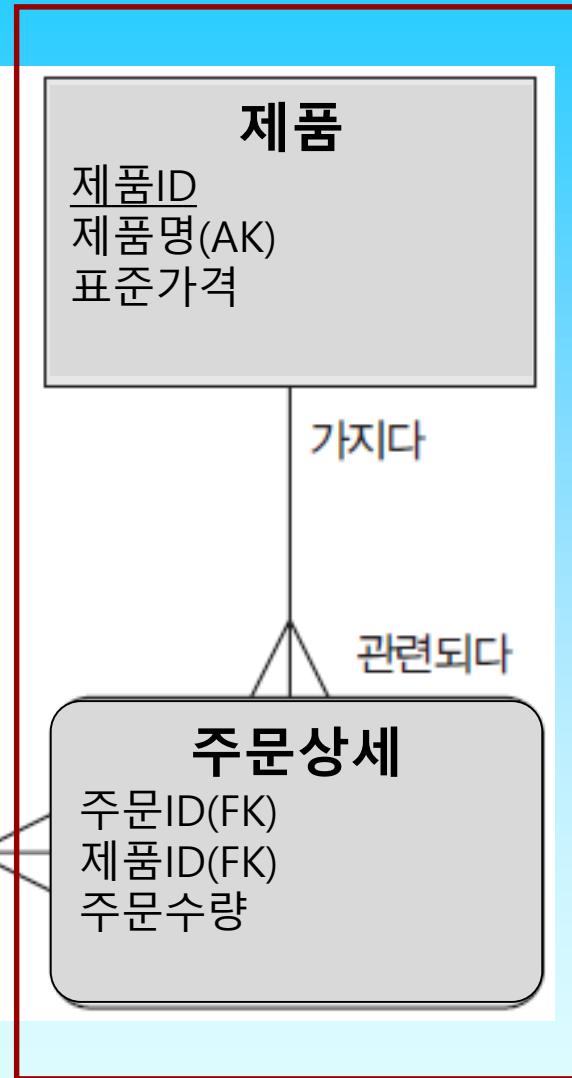
포함하다

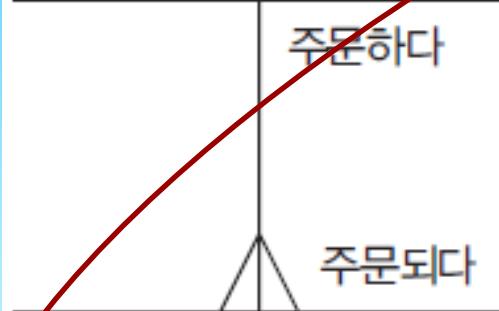
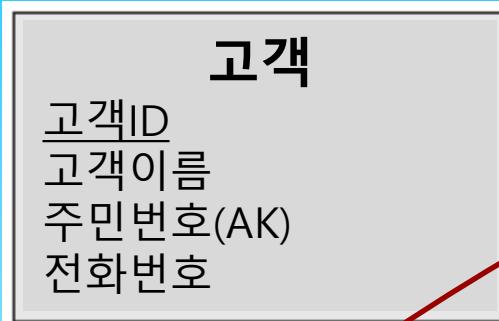
포함되다



하나의 제품은
여러 주문상세에
가질 수 있으며,
각 주문상세는
하나의 제품과
관련된다.

→ 일대다 관계





따라서, 하나의 주문은 여러 제품을 포함하고 한 제품은 여러 주문에 포함된다.
 → 다대다 관계

포함하다

포함되다

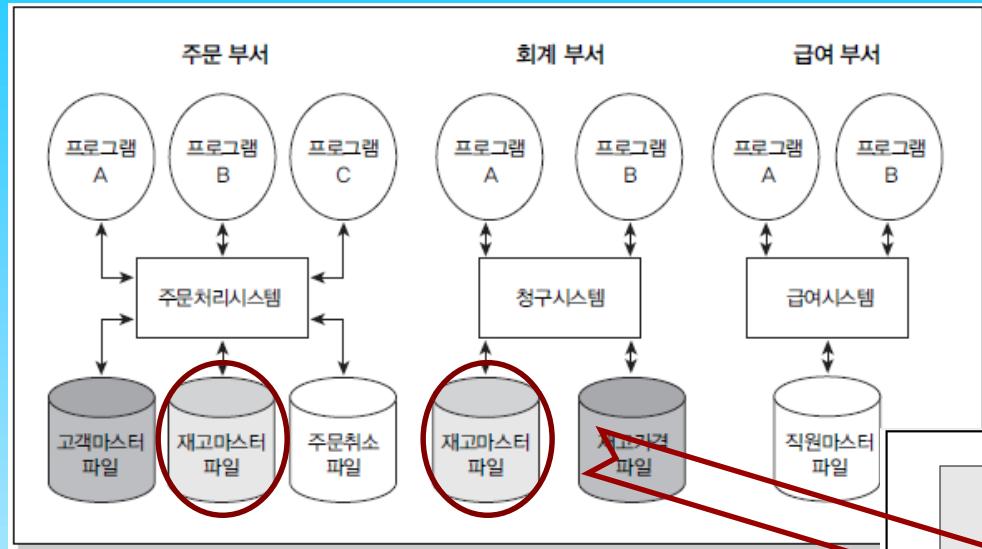


그림 1-3 Pine Valley Furniture Company의
오래된 파일처리 시스템

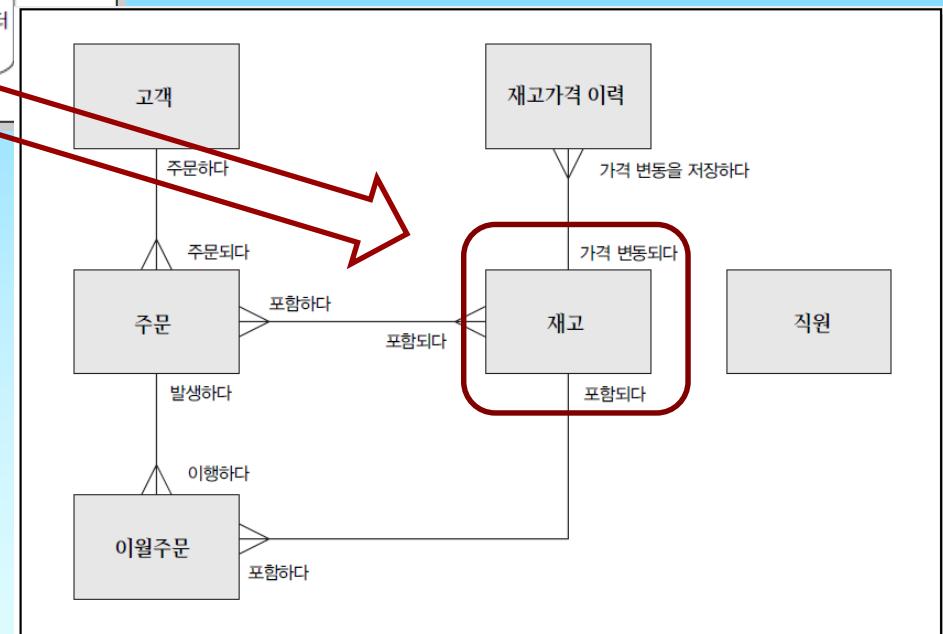
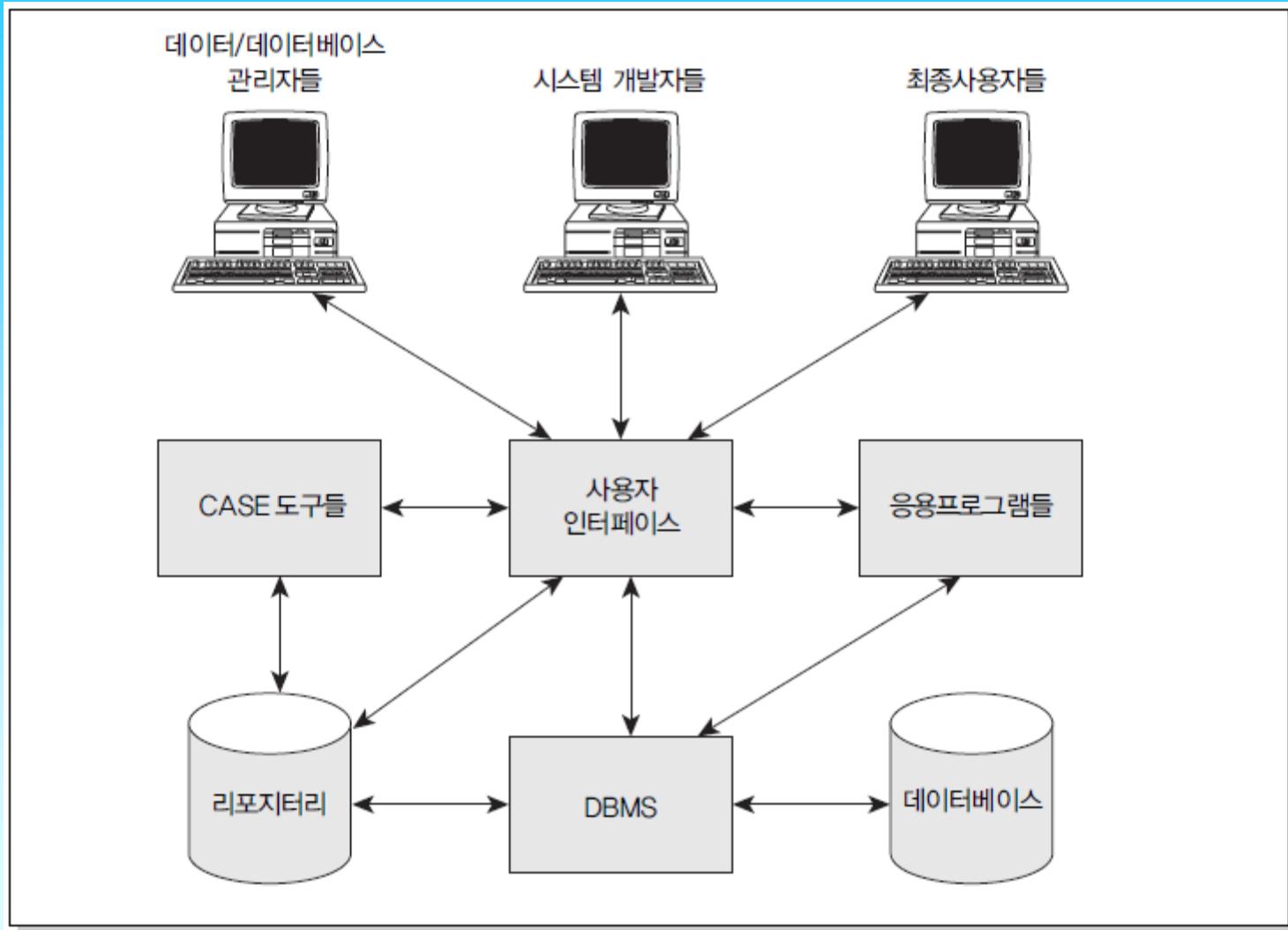


그림 1-4 [그림 1-3]의 일부분에 대한 전사적 데이터 모델

그림 1-5 데이터베이스 환경의 요소들



데이터베이스 환경의 요소들

- CASE도구들: computer-aided software engineering
- 리포지터리: 중앙집중화된 메타데이터 집합
- 데이터베이스관리시스템(DBMS) : 데이터베이스 관리를 위한 소프트웨어
- 데이터베이스: 데이터의 구조화된 집합
- 응용프로그램들: 데이터를 사용하는 소프트웨어
- 사용자 인터페이스: 사용자에게 표시하는 문자와 그림
- 데이터/데이터베이스 관리자: 데이터베이스 관리를 책임지는 사람
- 시스템 개발자들: 소프트웨어와 데이터베이스 설계를 책임지는 사람
- 최종사용자들: 데이터베이스와 응용프로그램을 사용하는 사람

데이터베이스 응용 범주

- 개인용 데이터베이스
- 워크그룹 데이터베이스
- 부서/부문 데이터베이스
- 전사적 데이터베이스
 - 전사적자원관리시스템(enterprise resource planning systems)
 - 데이터웨어하우스 구현

표 1-6 데이터베이스 응용 범주 요약

데이터베이스 유형	전형적인 사용자 수	전형적인 아키텍처	전형적인 데이터베이스 용량
개인용	1	데스크탑/휴대용 컴퓨터, PDA	메가바이트
워크그룹	5~25	클라이언트/서버(2-계층)	메가바이트~기ガ바이트
부서	25~100	클라이언트/서버(3-계층)	기ガ바이트
전사적	> 100	클라이언트/서버 (분산 또는 병렬 서버)	기ガ바이트~테라바이트
웹 연동	> 1000	웹 서버와 애플리케이션 서버들	메가바이트~기ガ바이트

그림 1-6
개인용 데이터베이스의
전형적인 데이터들

고객

Edit Robert Watson

General Note

Name & Contact Info

Name:

Last Name: Watson
First Name: Robert
Company: Automotive
Title: Custom Solutions

Contact Info:

Work • 813-821-1234
Home •
Fax • 813-821-5678
Other • 813-353-9012
E-Mail • rwatson@acs.com

Select the primary contact: Work

Chat & Web:

M •
M •
Website

Work Home Other

Address • 4500 W Kennedy Blvd
City: Tampa
State: FL
Zip: 33609
Country:

Additional Info

Custom:

Custom 1:
Custom 2:
Custom 3:
Custom 4:

Birthday: [] (ex: 4/1/2007)

Enable reminder on handheld: Days

OK Cancel New Help

Contact Photo: 

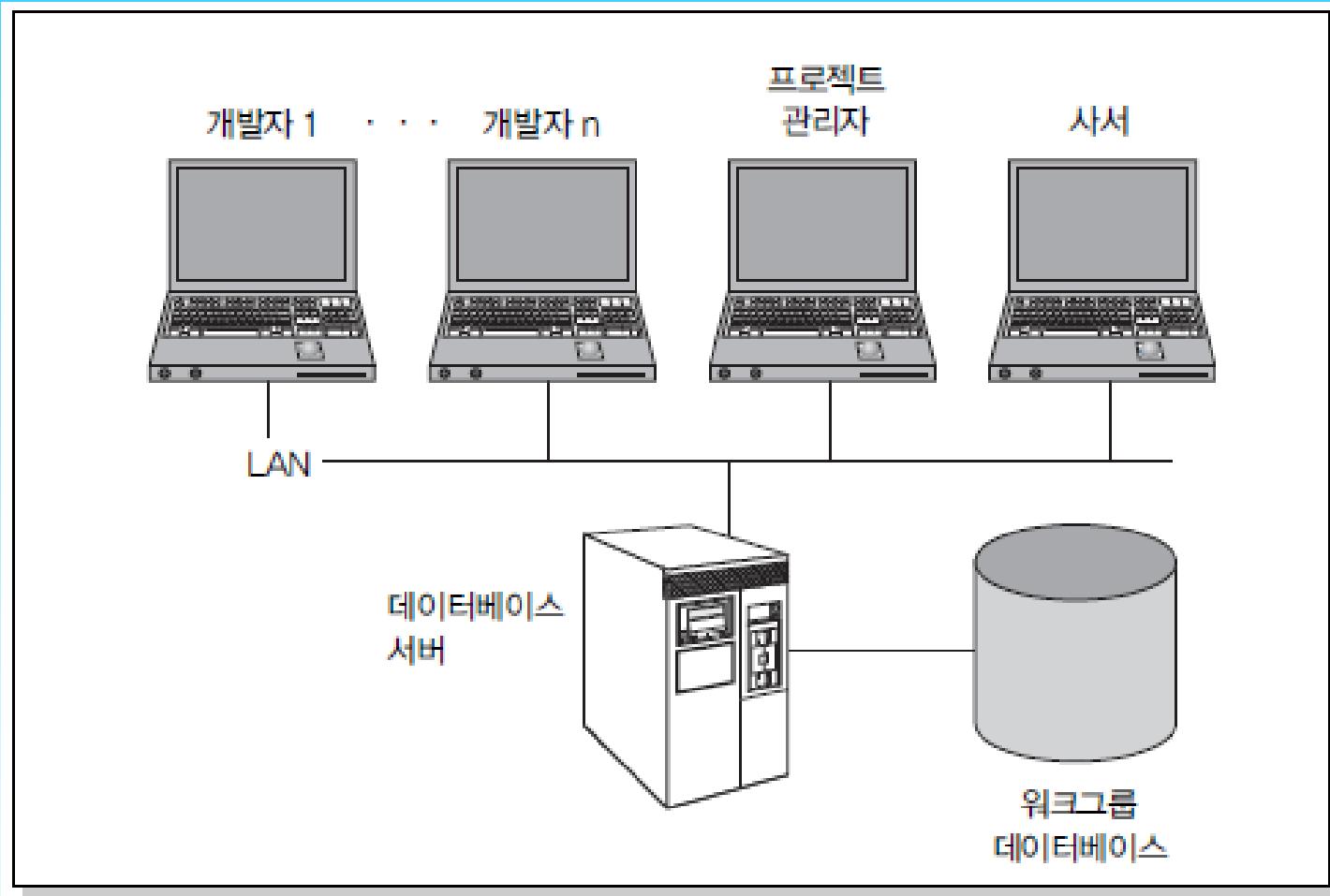
Category: Business

Private

고객 접촉 기록

Date	Time	Contact	Comments
08/04/2008	10:00AM	Roberts	Review proposal
08/19/2008	08:00AM	Roberts	Revise schedule
09/10/2008	09:00AM	Pearson	Sign contract
09/21/2008	02:00PM	Roberts	Follow up

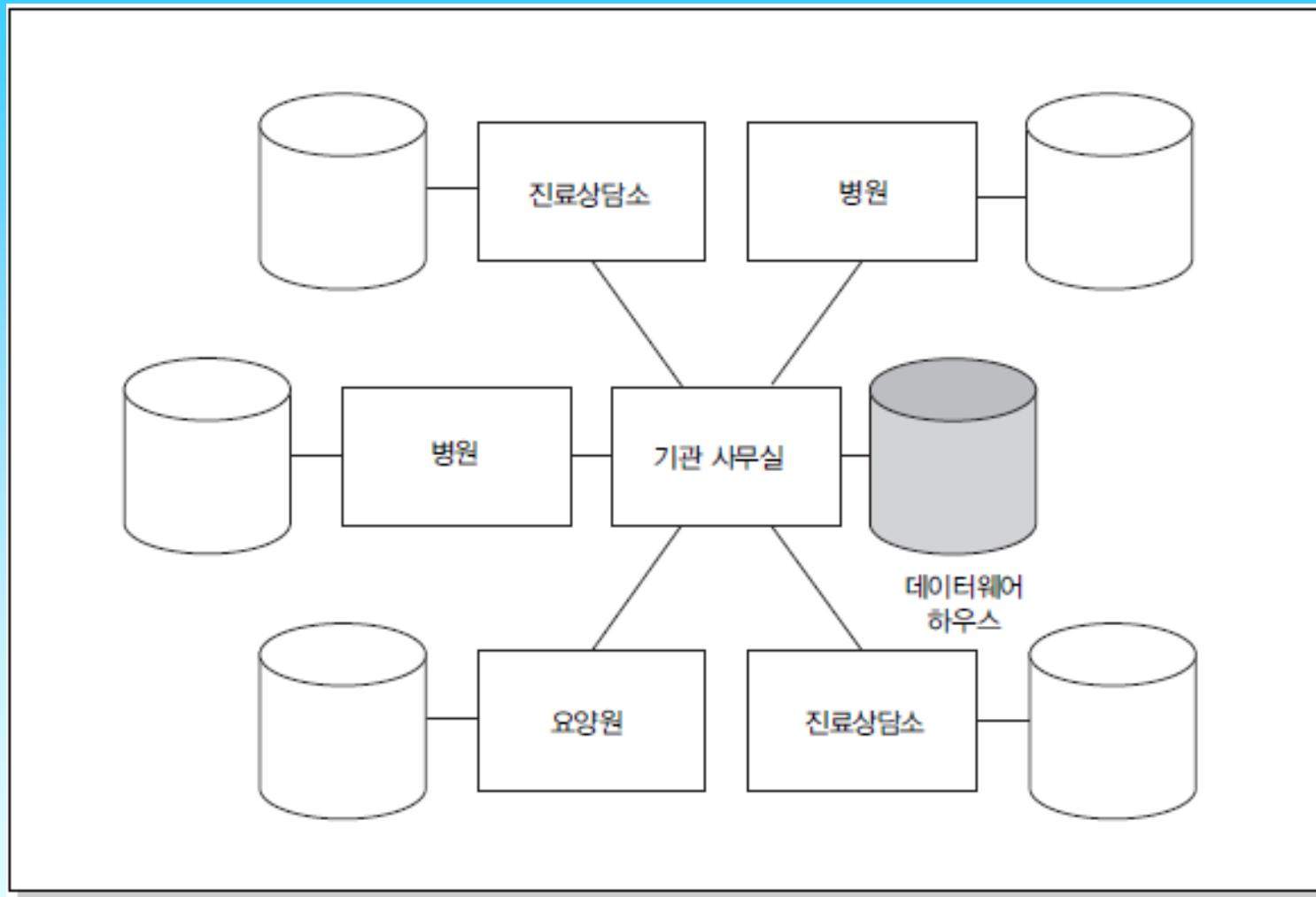
그림 1-7 LAN에 연결된 워크그룹 데이터베이스



전사적 데이터베이스 응용

- 전사적자원관리 (ERP)
 - 전체적인 영역을 통합 (제조, 금융, 영업, 마케팅, 재고, 회계, 인적 자원)
- 데이터웨어하우스
 - 다양한 운영 데이터베이스에서 파생된 통합 의사결정지원시스템

그림 1-8 전사적 데이터웨어하우스



웹 연동 데이터베이스

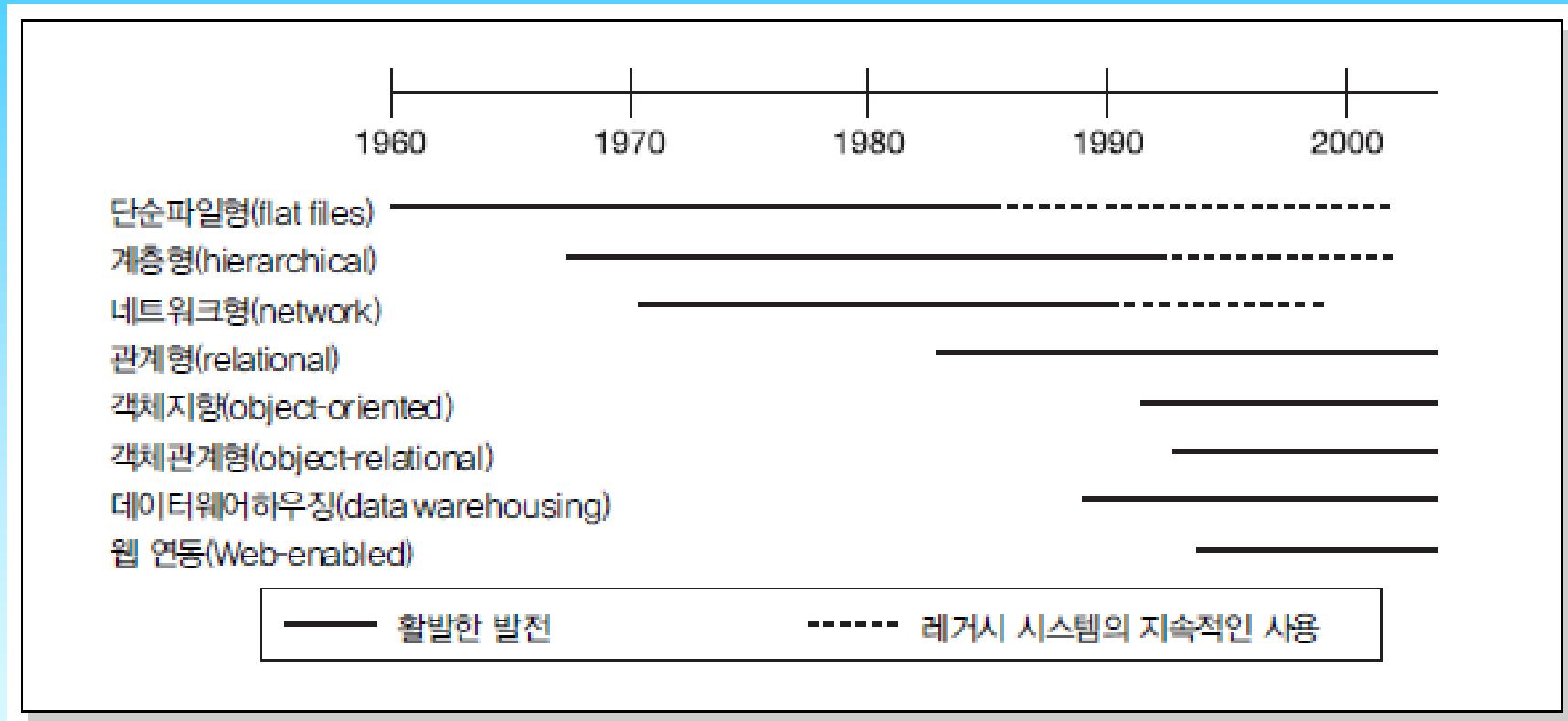
- 데이터베이스를 필요로 하는 웹 어플리케이션
 - 고객관계관리시스템 (CRM)
 - 기업과 소비자와의 거래 (B2C)
 - 전자문서교환 (EDI)
 - 사설 인트라넷 (Private intranets)
 - XML-defined Web services

웹 연동 데이터베이스 (cont.)

■ 제기되는 질문들

- 어떤 기술들이 사용되어야 하는가?
- 보안/프라이버시 보호
- 인터넷 기반의 거래를 통해 생성된 엄청난 양의 데이터들을 어떻게 관리해야 하는가?
- 데이터 품질을 어떻게 유지해야 하는가?

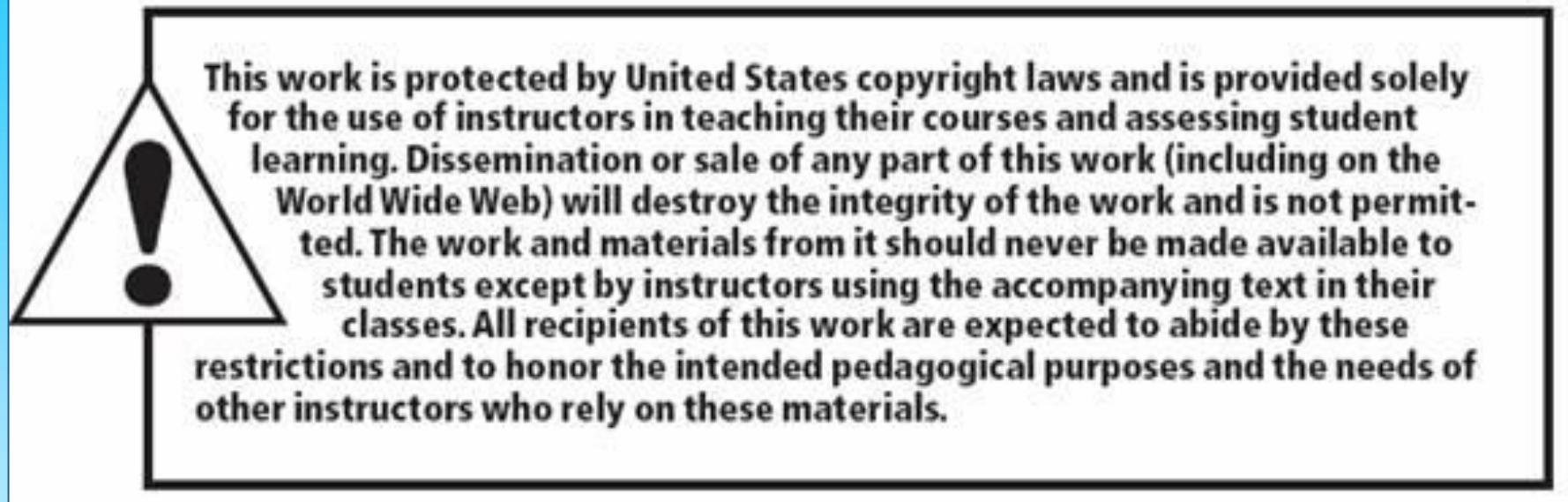
그림 1-9 데이터베이스 기술의 발전



HTML

```
<html>
  <head>
    <title>부산대학교 데이터베이스관리</title>
  </head>

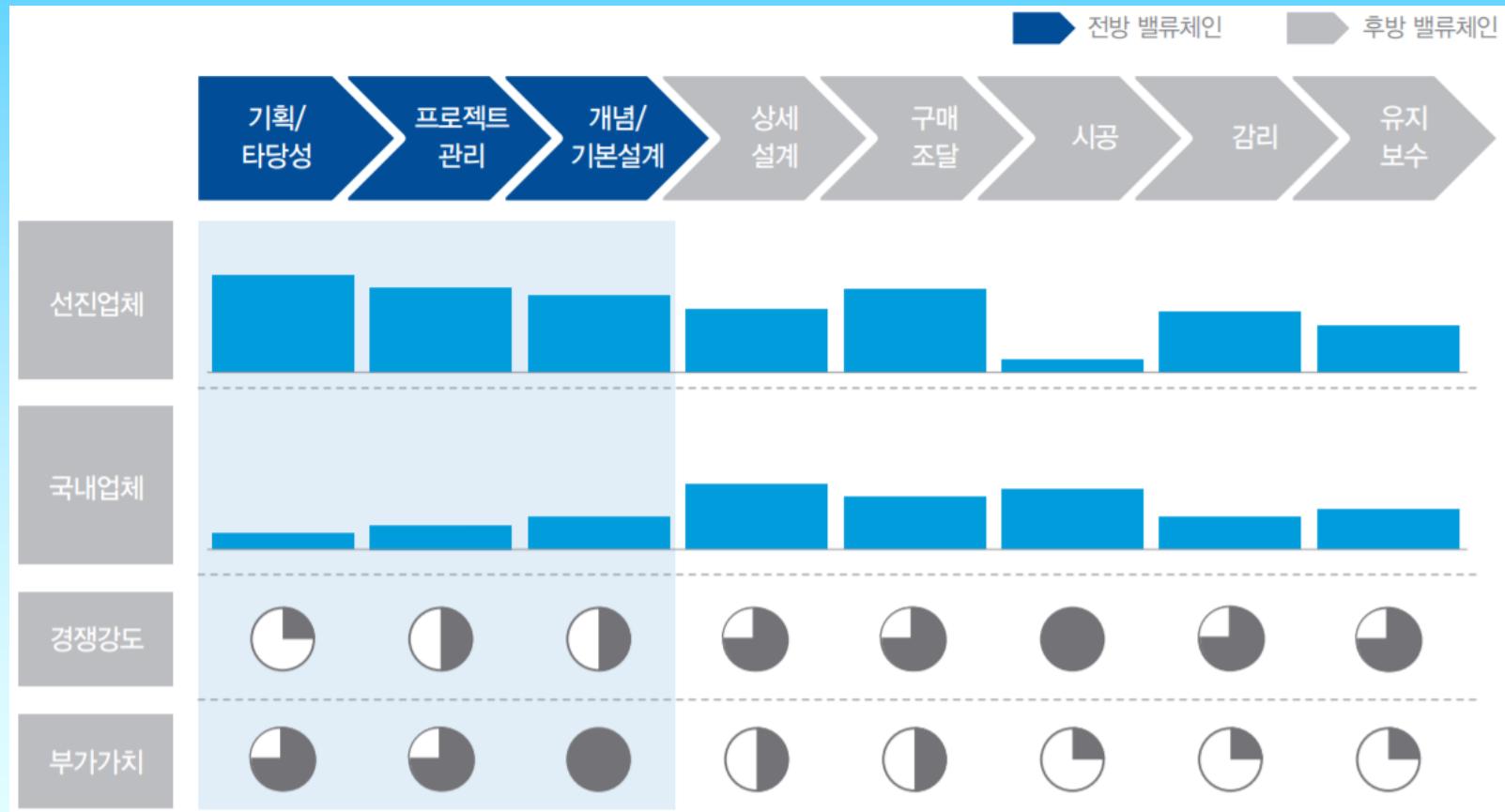
  <body>
    <h1>Database Management</h1>
    <p>이 자료는 <a href="http://plms.pusan.ac.kr">데이터베이 관리 수업</a>에서 배포합니다.</p>
    
      <font size="4">내가 제일 잘나가</font><br>
      <font color="red"> 내가 제일 잘나가 </font><br>
      <font face="돋움"> 내가 제일 잘나가 </font><br>
      <font size="4"><b> 내가 제일 잘나가 </b></font><br>
      <font color="red"><i> 내가 제일 잘나가 </i></font><br>
      <font face="돋움"><u> 내가 제일 잘나가 </u></font><br>
  </body>
</html>
```



All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. Printed in the United States of America.

Copyright © 2016 Pearson Education, Inc. Publishing as Prentice Hall

건설산업 전,후방 밸류체인 경쟁력



자료: 산업자원통산부