

김현승, 「디지털 스토리텔링」, 한국전통문화대학교 수업 자료, 2024.

데이터베이스 모델링

김현승

한국학중앙연구원 디지털인문학연구소

2024.04.30.



01

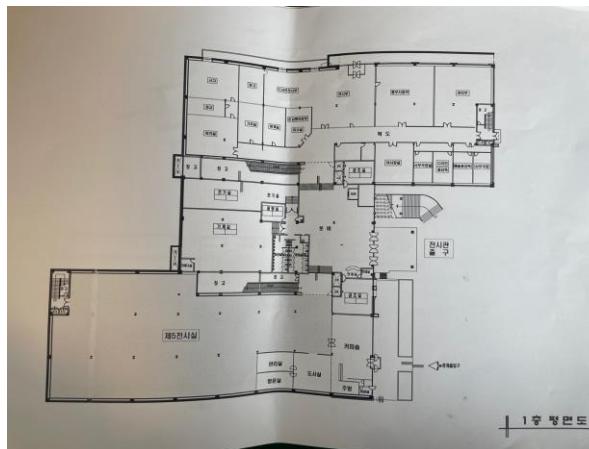
Part 1 데이터베이스 모델링

Contents

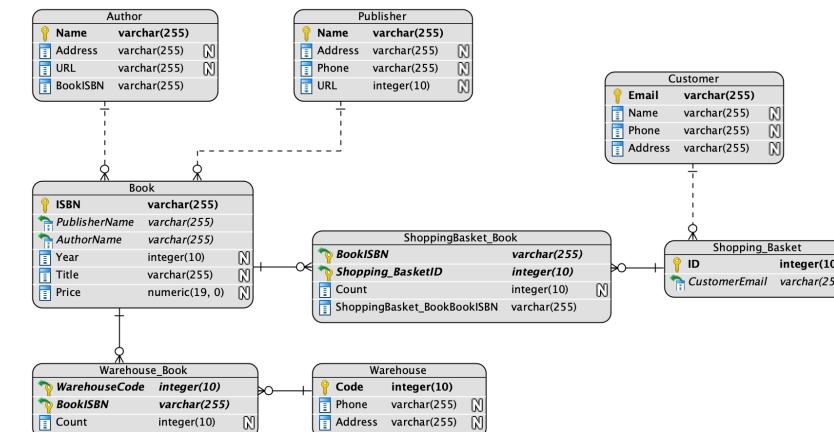
데이터베이스 모델링

데이터베이스 모델링 (Database Modeling)

데이터베이스를 설계할 때 효율적으로 데이터를 저장할 곳을 마련하기 위해 미리 설계하는 단계로, 테이블의 구조를 결정하는 과정이다.



광주 비엔날레 설계도



관계형 데이터베이스 ERD

Contents

데이터베이스 모델링



데이터베이스 모델링의 필요성

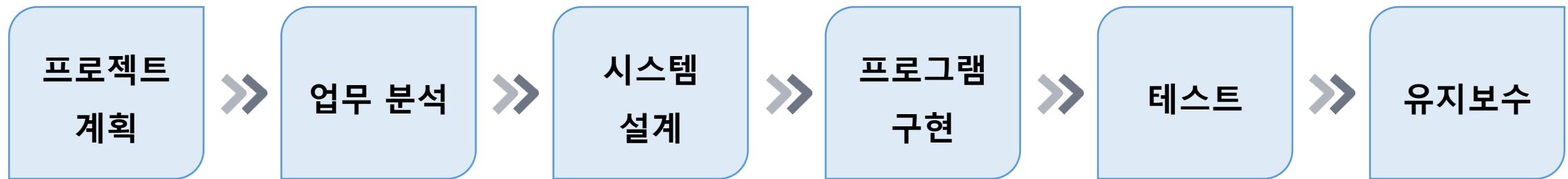
- DBMS 구축에 필요한 다양한 기술을 효율적으로 적용하는 방안을 제시한다.
- 데이터베이스 설계 및 생성 속도와 효율성을 촉진시킨다.
- 조직의 데이터를 문서화하고 데이터 관련 시스템을 설계할 때 일관성을 조정한다.
- 업무 조직과 기술 조직 간의 의사소통을 원활히 하는 도구 또는 중재의 역할을 한다.

Contents

데이터베이스 모델링



프로젝트 개발 단계

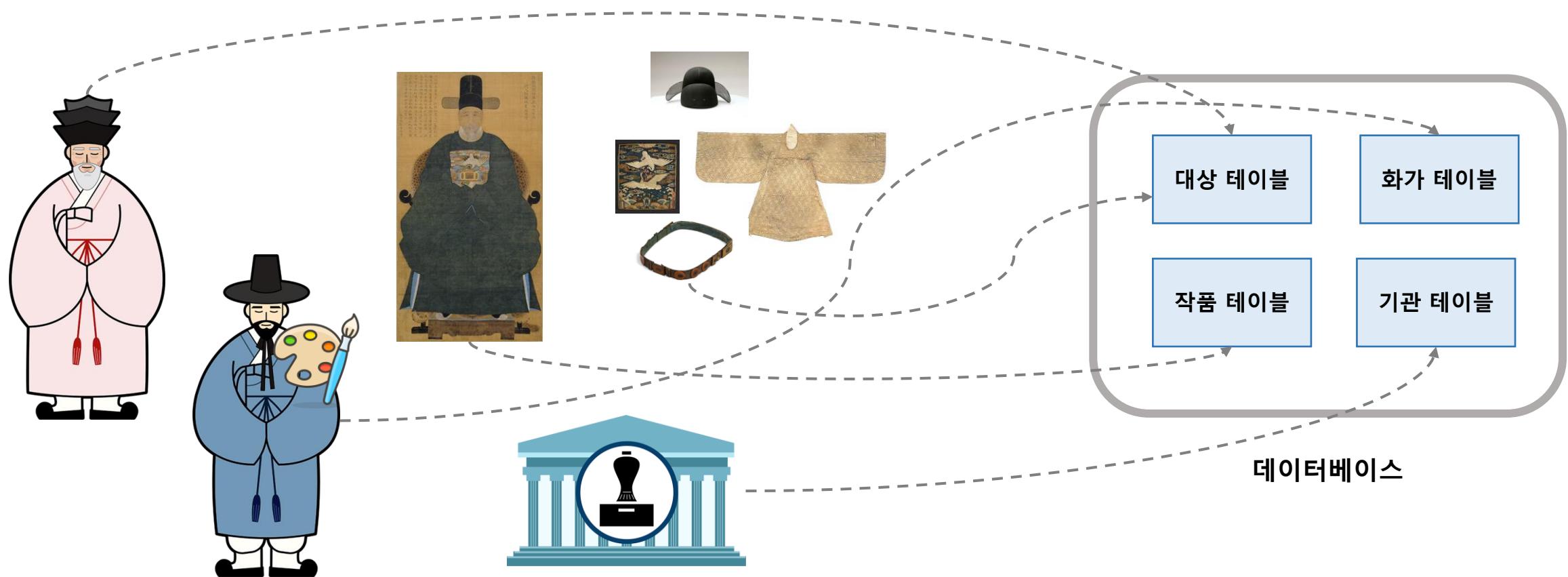


Contents

데이터베이스 모델링

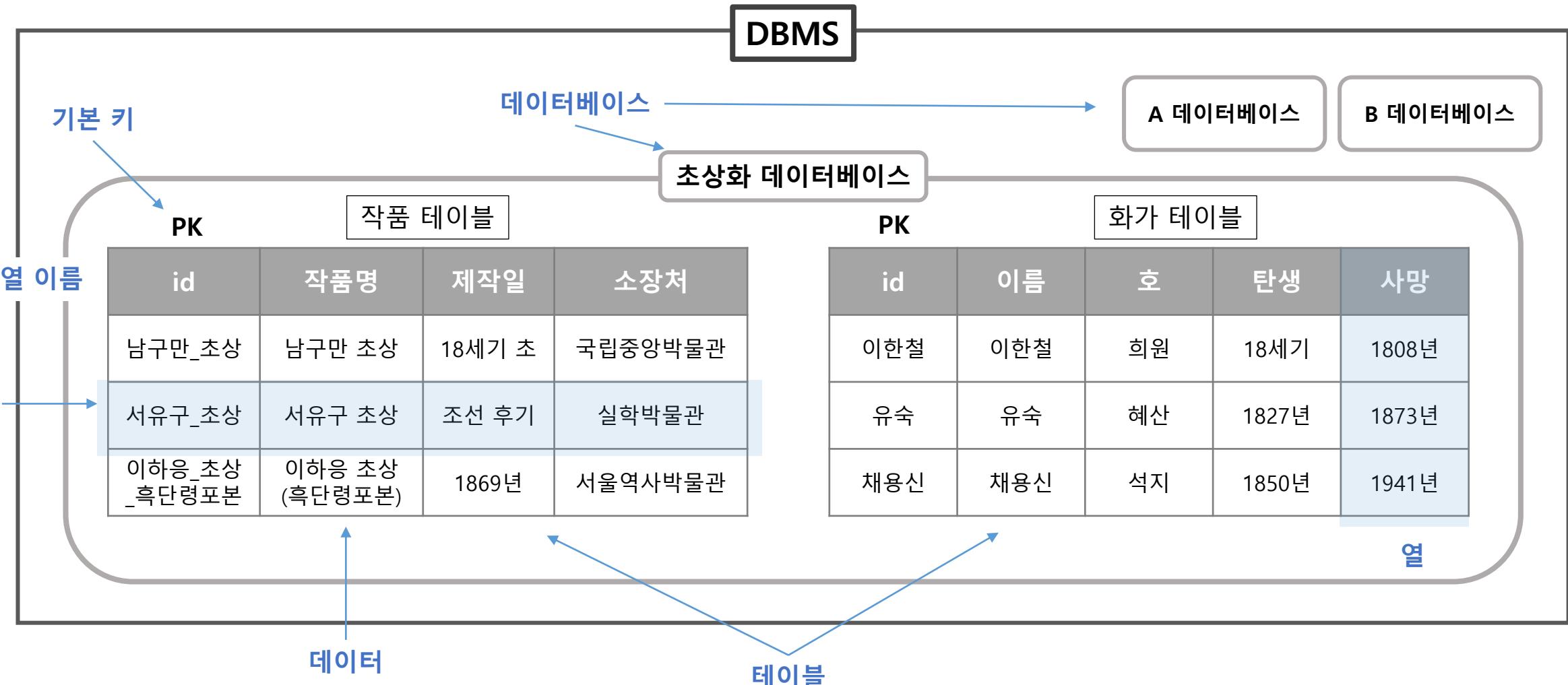


데이터베이스 모델링 구성



Contents

데이터베이스 모델링



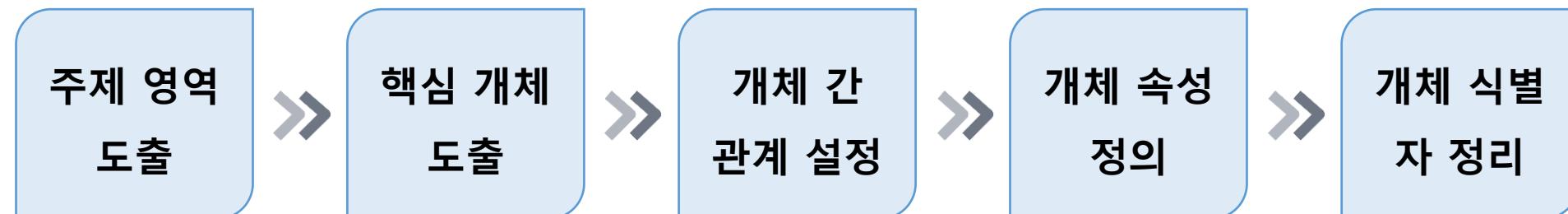
Contents

데이터베이스 모델링



데이터베이스 모델링의 유형: 개념적 데이터 모델

- 프로젝트 관련 비지니스 이해 관계자와 분석가가 개념적으로 모델을 생성하는 것.
- 공식 데이터를 활용하여 모델링하는 것보다는 요구 사항을 도출하고, 프로젝트의 범위와 설계를 어떻게 할 것인지를 정의하는 단계.



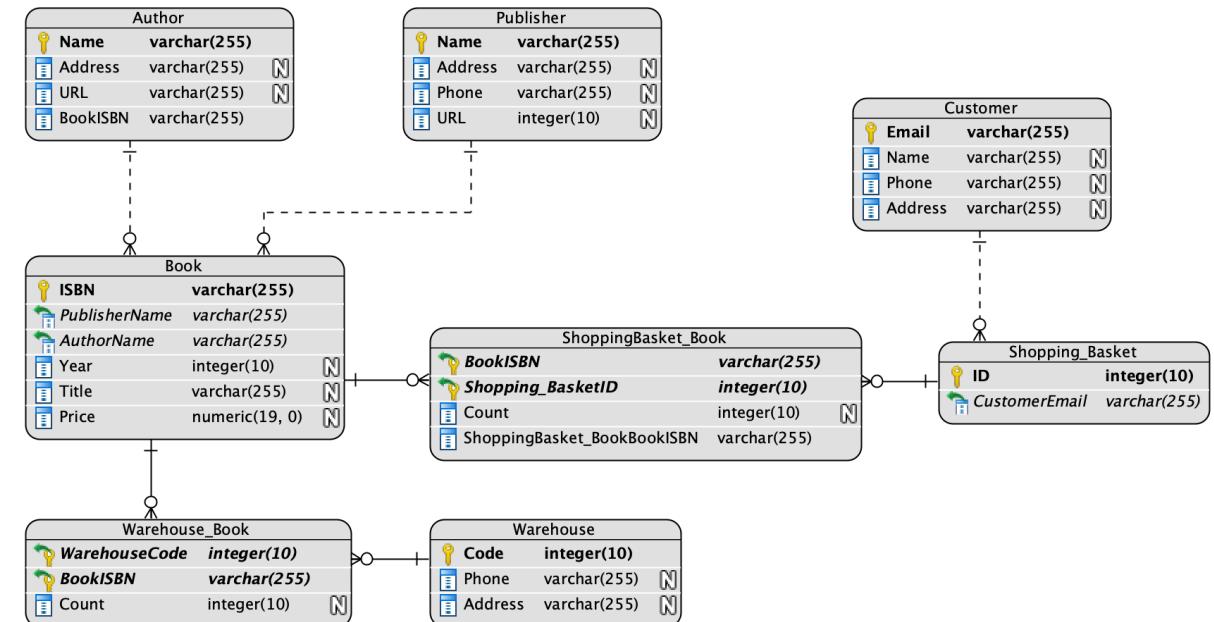
Contents

데이터베이스 모델링

데이터베이스 모델링의 유형:

논리적 데이터 모델

- 개념적 데이터 모델에서 도출한 엔티티를 기술적 데이터 구조와 연결하는 단계.
- 주로 데이터 아키텍트와 분석가 등이 협업하여 제작.
- 공식 데이터 모델링 시스템 중 하나를 선택해 모델링 작성. 논리적 ERD.



Contents

데이터베이스 모델링



데이터베이스 모델링의 유형: 물리적 데이터 모델

- 논리적 데이터 모델을 관련자들이 정한 DBMS 기술에 접목하고 해당 DB 언어를 사용해서 만든다.
- 논리적 데이터 모델과 최종적으로 선택할 DBMS 사이에서 다리 역할을 한다.

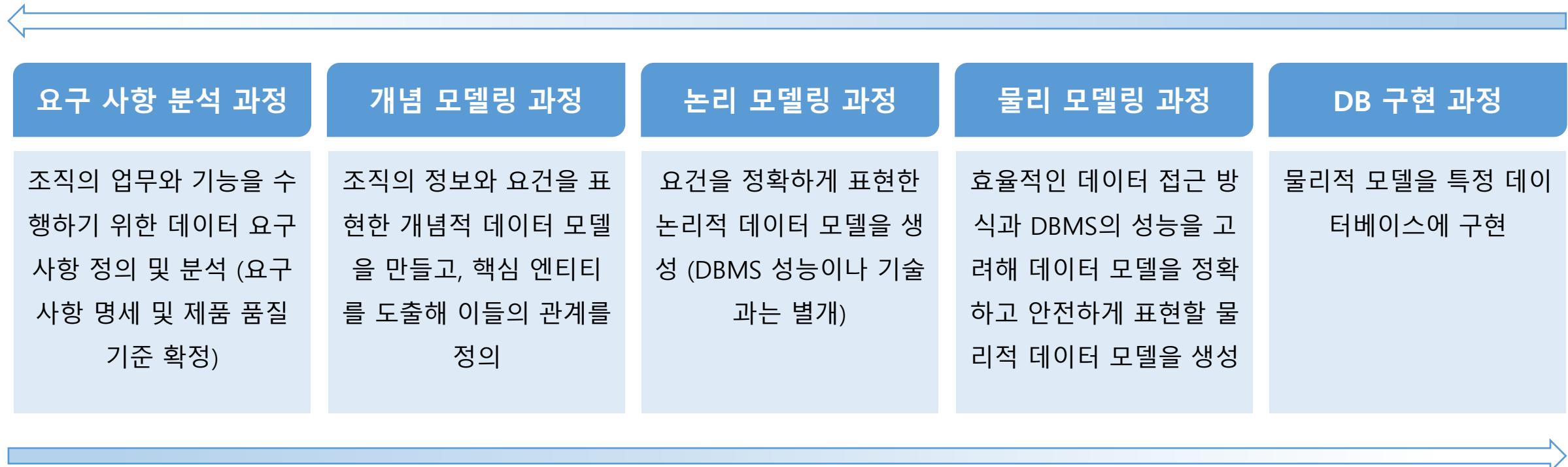
Contents

데이터베이스 모델링



데이터베이스 모델링 단계

추상화, 단순화



구체화, 세분화



Part 2

데이터 타입(type)

Contents

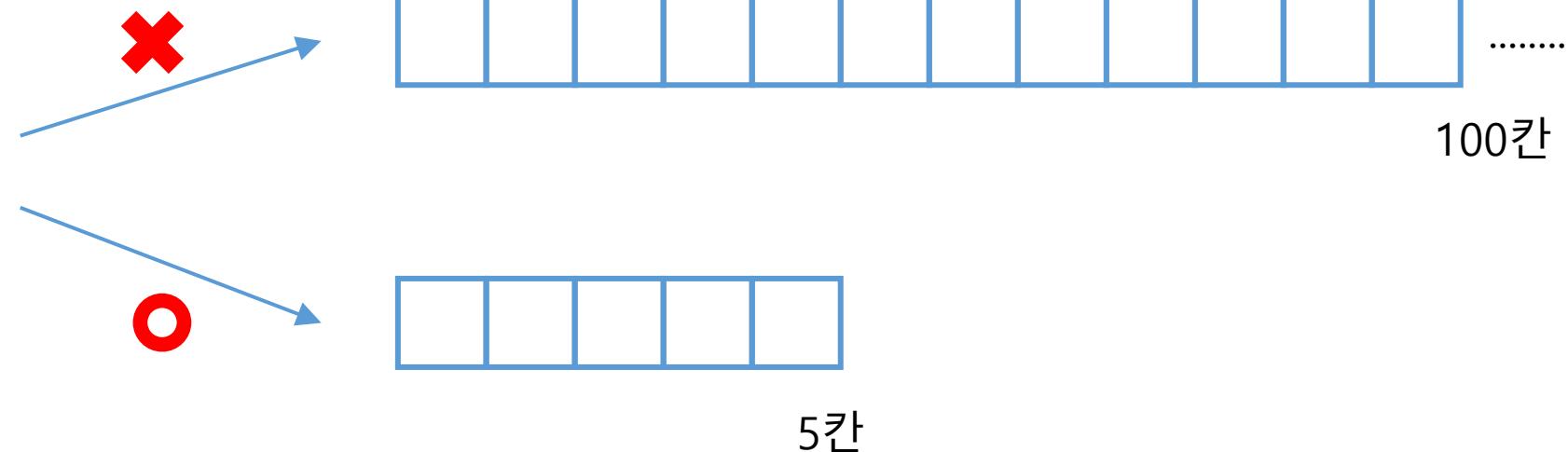
데이터 타입



데이터 타입 : 숫자형 / 문자형 / 날짜형 ...

김현승

사람 이름
(Data)



Contents

데이터 타입 : 숫자형



숫자형은 정수, 실수 등의 숫자로 된 데이터 유형을 말한다.

정수

12312348

실수

3.151592

시간

1753/1/1

Contents

데이터 타입 : 정수형



정수형은 소수점이 없는 숫자, 즉 인원 수, 가격, 수량 등에 많이 사용한다.

데이터 타입	데이터 크기(byte)	숫자 범위(signed)	숫자 범위(unsigned)
TINYINT	1	-128 ~ 127	0 ~ 255
SMALINT	2	-32,768 ~ 32,767	0 ~ 65535
MEDINUMINT	3	-8,388,608 ~ 8,388,607	0 ~ 16,777,215
INT	4	-2^{31} (약 -21억) ~ $2^{31}-1$ (약 21억)	$0 \sim 2^{32}-1$ (약 42억)
BIGINT	8	-2^{63} (약 -900경) ~ $2^{63}-1$ (약 900경)	$0 \sim 2^{64}-1$

Contents

데이터 타입 : 실수형



실수형은 소수점이 있는 숫자를 저장할 때 사용한다.

데이터 타입	데이터 크기(byte)	설명
FLOAT	4	소수점 아래 7자리까지 표현
DOUBLE	8	소수점 아래 15자리까지 표현
DECIMAL	5 ~ 17	소수를 저장하지만 내부적으로는 문자 형태로 저장되는 타입
NUMERIC	5 ~ 17	DECIMAL과 동의어

Contents

데이터 타입



숫자

12312348

문자

가야sTS高

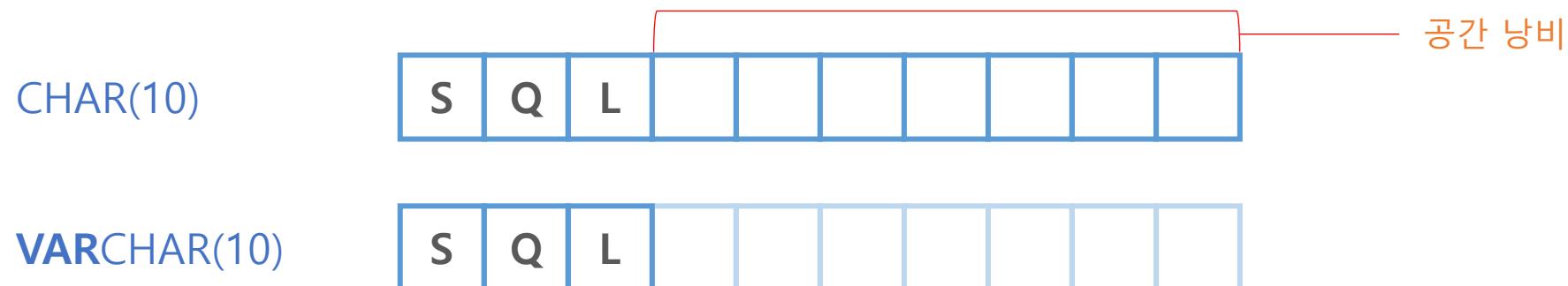
Contents

데이터 타입 : 문자형



문자형은 다양한 문자를 저장할 수 있는 데이터 유형이다. 입력할 최대 글자의 개수를 지정해야 한다.

- 고정 길이 : 실제 값을 입력하지 않아도 지정한 만큼 저장 공간을 사용
- 가변 길이 : 실제 입력한 값의 크기만큼만 저장 공간을 사용



Contents

데이터 타입 : 문자형



데이터 타입	데이터 크기 (byte)	설명
CHAR	1~255	char(0) 도 지원하며, 실제값이 M보다 작을 때 남는 자릿수만큼 공백을 붙여서 저장한다. 1부터 최대 255의 자릿수를 지원하며, 지정한 용량만큼 바이트를 사용 한다.
VARCHAR	1~16383	지정할 수 있는 길이는 1부터 255까지이며, 지정한 길이보다 작은 데이터를 저장 할 때 필요한 길이만큼 저장된다. CHAR 보다 기억장치를 효율적으로 사용할 수 있다. 지정한 용량 + 1바이트 사용.
NVARCHAR	1~4000	가변길이 문자열이며, 유니코드를 사용할 수 있다.
BINARY	1~255	고정 길이 문자열로, 0~255 크기의 문자열까지 저장 가능하다.
VARBINARY	1~65535	가변 길이 문자열이다.

Contents

데이터 타입



유니코드
unicode

nvarchar

English

한글

漢字

Contents

데이터 타입



전화번호

010-3434-5678

CHAR

VARCHAR

Contents

데이터 타입 : 대량의 데이터 형식



데이터 타입		데이터 크기(byte)	설명
TEXT 형식	TEXT	1~165535	최대 길이가 65,535byte인 문자열 데이터 저장
	LONGTEXT	1~4294967295	최대 길이가 4,294,967,295byte인 문자열 데이터 저장
BLOB 형식	BLOB	1~65535	최대 길이가 65,535byte인 데이터 저장
	LONGBLOB	1~4294967295	최대 길이가 4,294,967,295byte인 문자열 데이터 저장

Contents

데이터 타입 : 날짜형 및 시간형



날짜형 / 시간형 / 날짜 및 시간형 은 날짜 또는 시간만 저장하거나 날짜와 시간을 함께 저장하는 데이터 유형이다.

데이터 타입	데이터 크기 (byte)	설명
TIME	3	HH:MM:SS(시:분:초) 형태의 데이터에 사용하고, 범위는 -838:59:59~838:59:59 이다.
DATE	3	날짜만 있는 데이터에 사용하고, YYYY-MM-DD(연-월-일) 형태이다. 범위는 1000-01-01~9999-12-31 이다.
YEAR	1	연도만 저장하는 자료형이다.
DATETIME	8	날짜와 시간을 저장한다. YYYY-MM-DD HH:MM:SS 형식으로 사용한다.
TIMESTAMP	4	날짜와 시간 부분을 모두 포함하는 데이터에 사용한다. 1970-01-01 00:00:00 이후부터 초를 숫자로 저장하는 자료형이다.

References

참고문헌



- 김현 · 임영상 · 김바로, 『디지털 인문학 입문』, 한국외국어대학교 지식출판원, 2016. : https://dh.aks.ac.kr/Edu/wiki/index.php/디지털_인문학_입문
- 김현, 「인문정보 데이터베이스」, 한국학중앙연구원 한국학대학원 수업 자료, 2020.
- 김바로, 「인문정보 데이터베이스」, 한국학중앙연구원 한국학대학원 수업 자료, 2022.
- 한국학중앙연구원 한국학대학원 인문정보학과 SQL 스터디 자료, 2020.
- 강성욱, 『Do it! MySQL로 배우는 SQL 입문』, 이지스퍼블리싱, 2024.
- 우재남, 『혼자 공부하는 SQL』, 한빛미디어, 2021.