



# Data Modeling

안 화 수

# Data Modeling

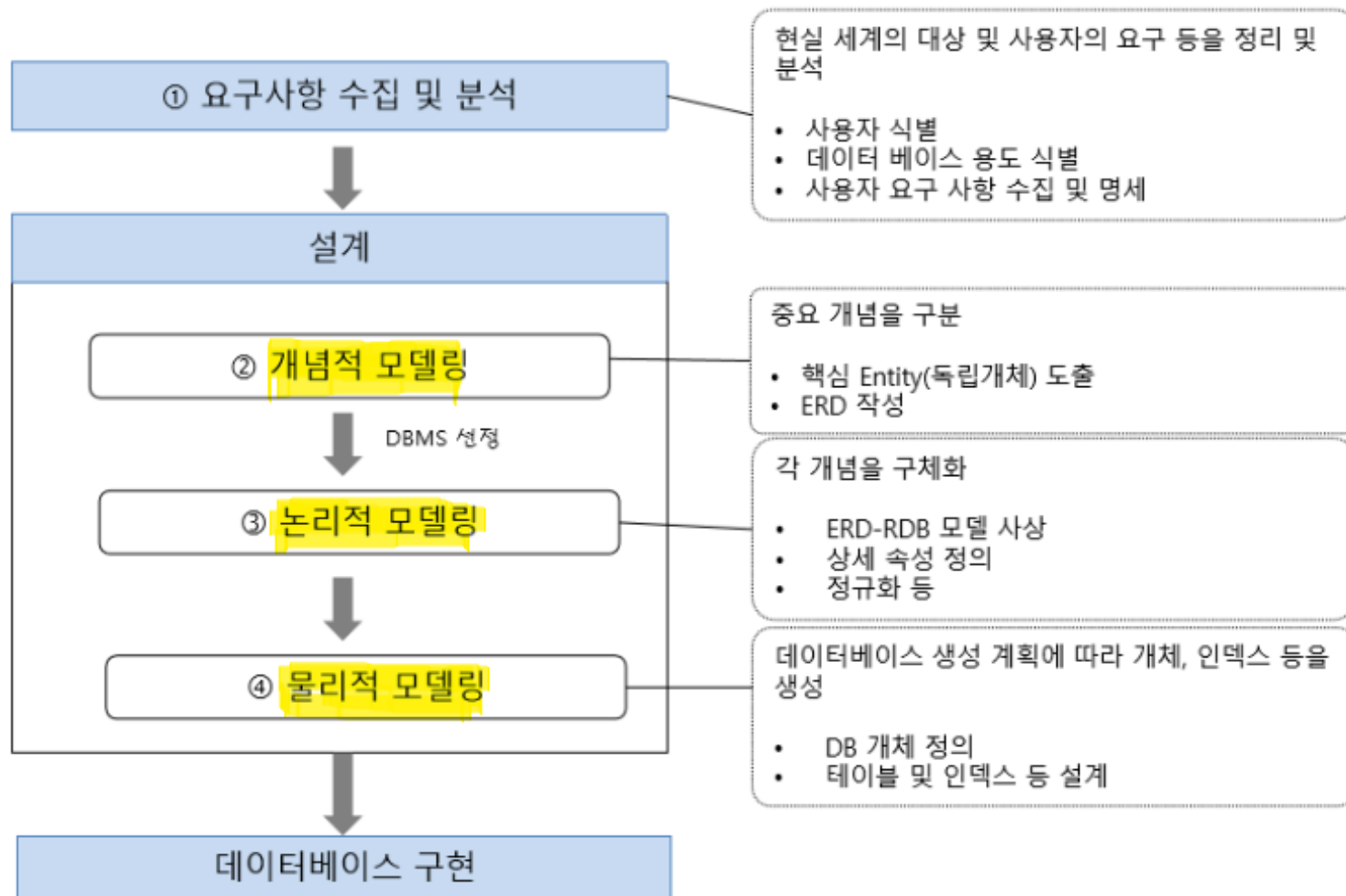
## ❖ Data Modeling 이란?

데이터 모델링은 데이터 흐름을 도식화하는 과정입니다. 데이터베이스 구조를 생성할 때 설계자는 데이터가 어떻게 데이터베이스로 유입되고 빠져나오는지에 관한 다이어그램으로 표현합니다. 이런 흐름에 관한 다이어그램은 데이터 형식, 구조, 데이터베이스 처리 기능의 특징을 정의하는데 사용되어 데이터 흐름 및 요구사항을 효율적으로 처리합니다.

# 데이터베이스 생명주기

- ❖ 요구사항 수집 및 분석
- ❖ 설계
- ❖ 구현
- ❖ 운영
- ❖ 감시 및 개선

# 데이터 모델링 과정



# 3가지 모델링

- ❖ 개념적 모델링

개체와 개체들 간의 관계에서 ER다이어그램을 만드는 과정

- ❖ 논리적 모델링

ER다이어그램을 사용하여 관계 스키마 모델을 만드는 과정

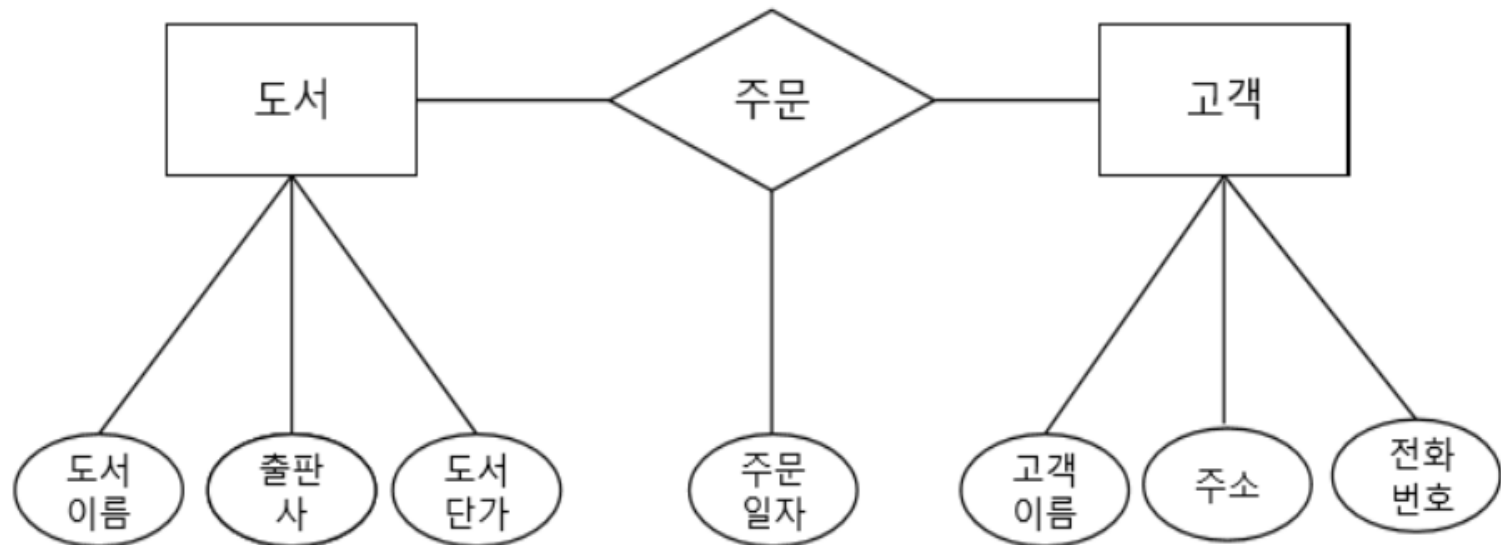
- ❖ 물리적 모델링

관계 스키마 모델의 물리적 구조를 정의하고 구현하는 과정

# 개념적 모델링

## ❖ 개념적 모델링이란?

1. 요구사항을 수집하고 분석한 결과를 토대로 업무의 핵심적인 개념을 구분하고 전체적인 뼈대를 만드는 과정입니다.
2. 개체(Entity)를 추출하고 개체들 간의 관계를 정의하여 **ER다이어그램을 만드는 과정**까지를 말합니다.



# 논리적 모델링

## ❖ 논리적 모델링이란?

1. 개념적 모델링에서 만든 ER 다이어그램을 사용하려는 DBMS에 맞게 사상(Mapping)하여 실제 데이터베이스로 구현하기 위한 **관계 스키마 모델을 만드는 과정**입니다.
2. 논리적 모델링을 하는 과정에는 상세속성 추출, 정규화 수행, 데이터 표준화 수행이 있습니다.



도서 (도서번호, 도서이름, 출판사이름, 도서단가)

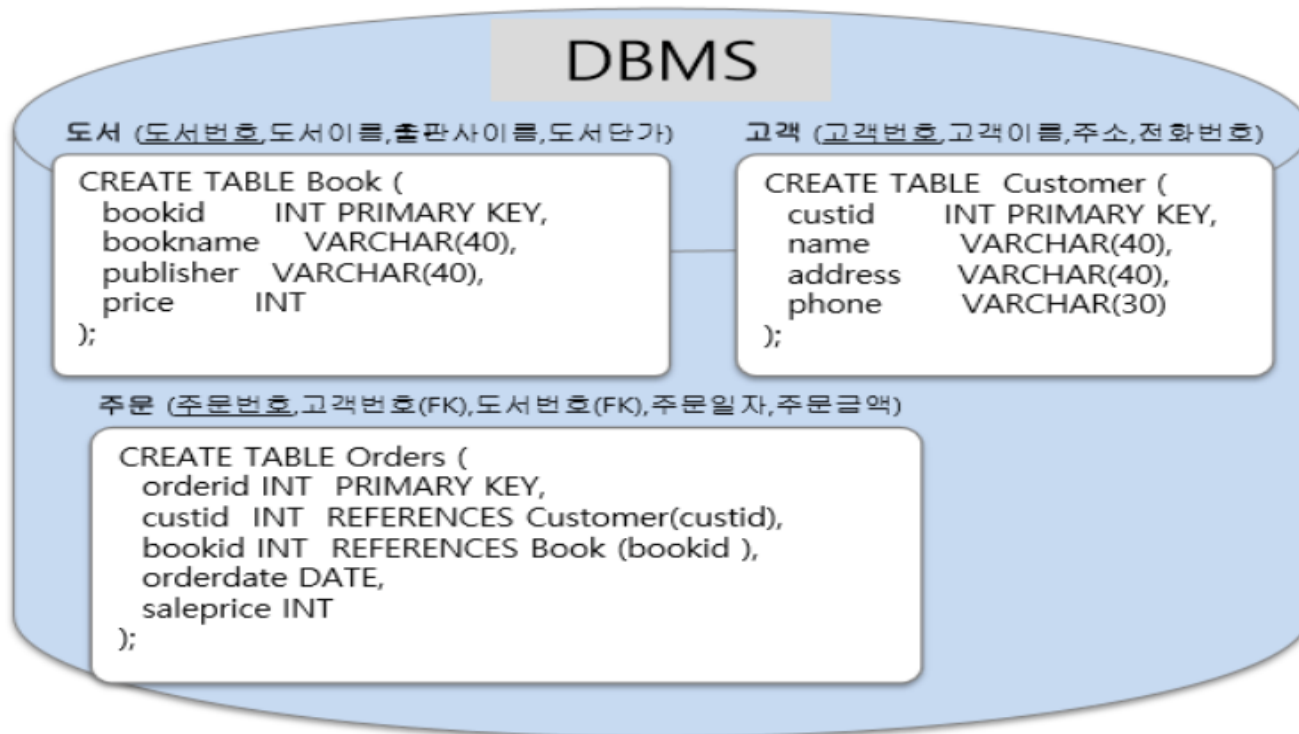
고객 (고객번호, 고객이름, 주소, 전화번호)

주문 (주문번호, 고객번호(FK), 도서번호(FK), 주문일자, 주문금액)

# 물리적 모델링

## ❖ 물리적 모델링이란?

물리적 모델링은 작성된 논리적 모델을 실제 컴퓨터의 저장 장치에 저장하기 위한 **물리적 구조를 정의하고 구현하는 과정**입니다.





# Modeling 제품

구분 \ 제품	ERWin	ER/Studio	DA#	eXERD
버전	7.X	9.X	3.2	2.X
실행환경	Windows	Windows	Windows	멀티플랫폼
공동작업	○	○	○	V2.0 이후
개념 모델링	X	○	○	X
논리/물리 통합 모델링	X	X	X	○
다중 다이어그램	○	○	○	○
관계선 자동배치	X	X	X	○
논리적 컬럼 그룹	X	X	○	○
용어/도메인	○	○	○	○
Forward/Reverse	○	○	○	○
다기종 DBMS 지원	○	○	○	△ (現 Oracle, MS SQL, DB2, MySQL)
사용자편의성	△	△	△	○
가격	고가	고가	고가	저가

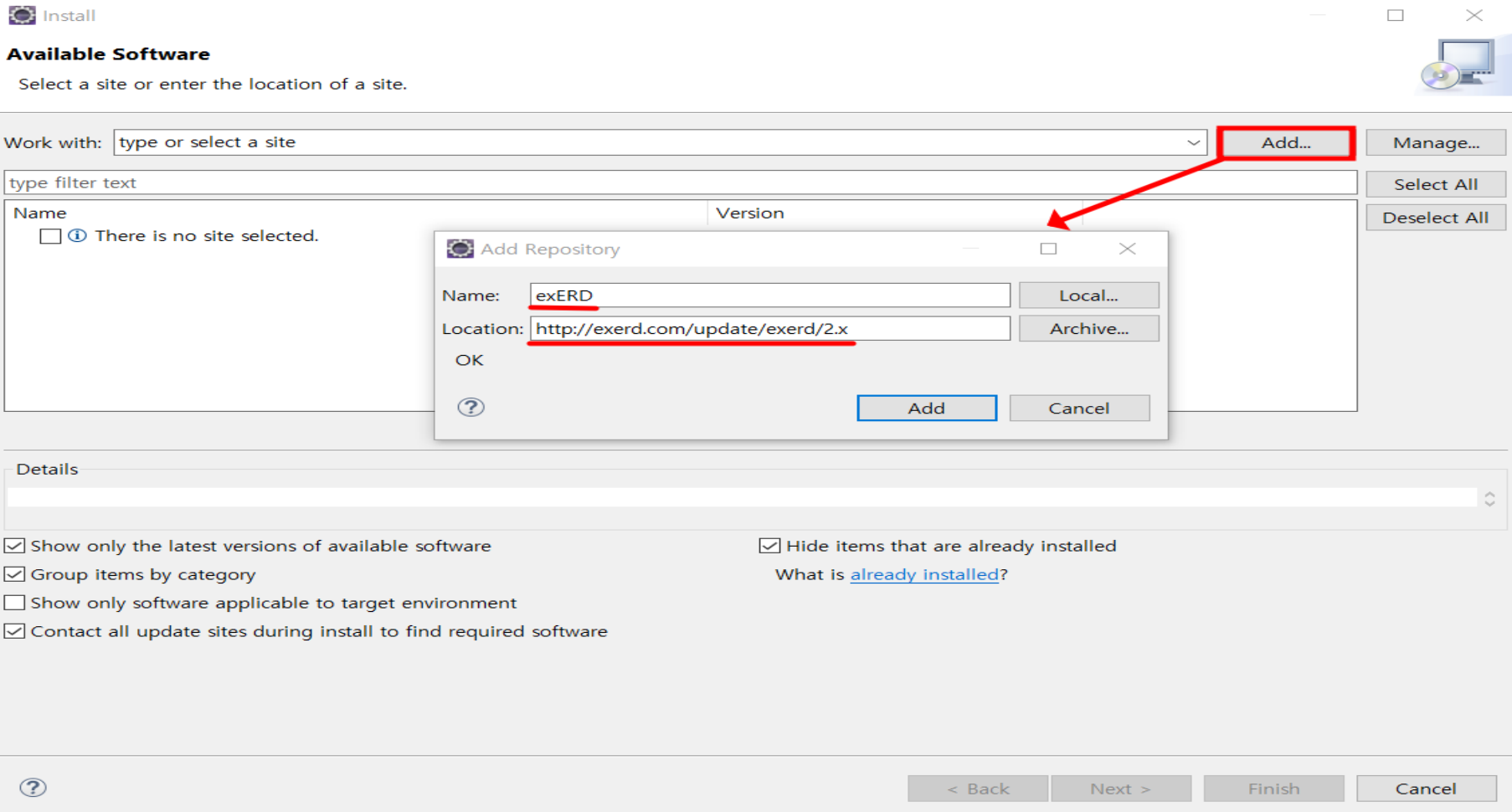
## ❖ exERD 특징

1. Oracle, MySQL, Microsoft SQL Server, IBM DB2 등 다양한 데이터베이스를 대상으로 포워드/리버스 엔지니어링 기능을 지원하고 있습니다.
2. 이클립스 기반의 지능형 E-R 모델링 도구입니다.
3. 이클립스 플러그인을 지원 합니다.

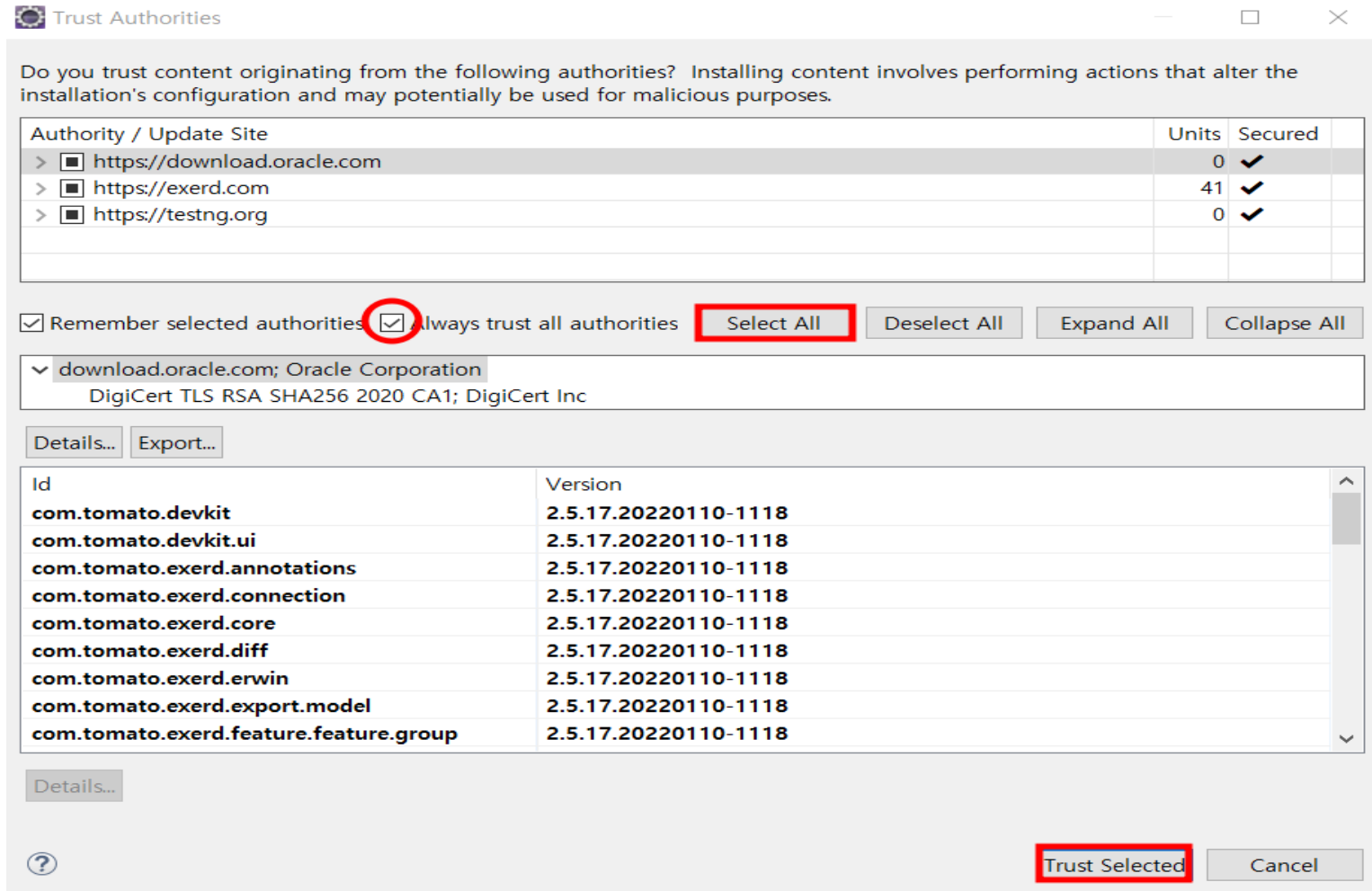
<http://exerd.com/update/exerd/2.x>

# exERD

❖ 이클립스에 exERD 플러그인 설치  
[Help] – Install New Software... 클릭

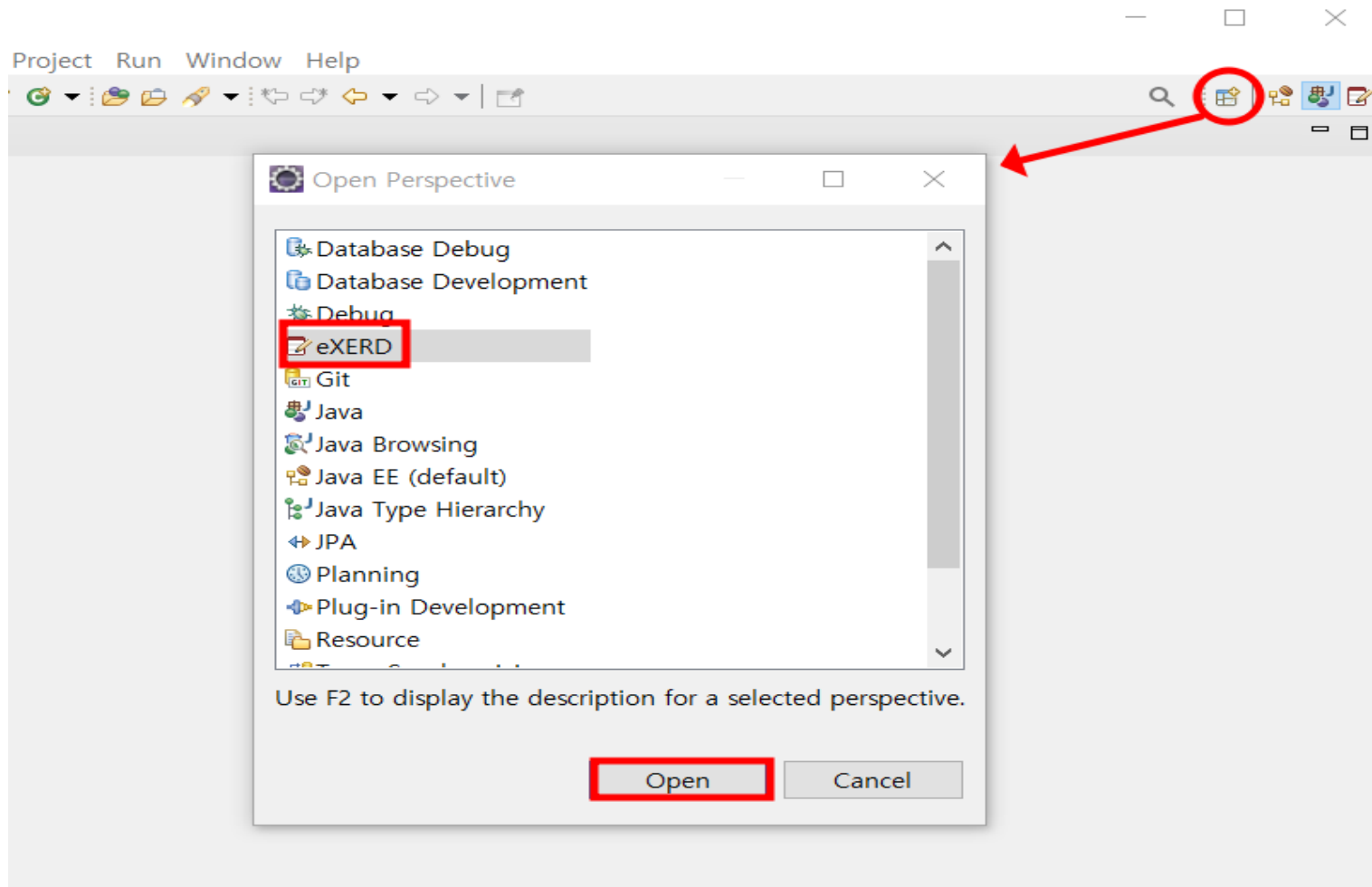


## ❖ 이클립스에 exERD 플러그인 설치

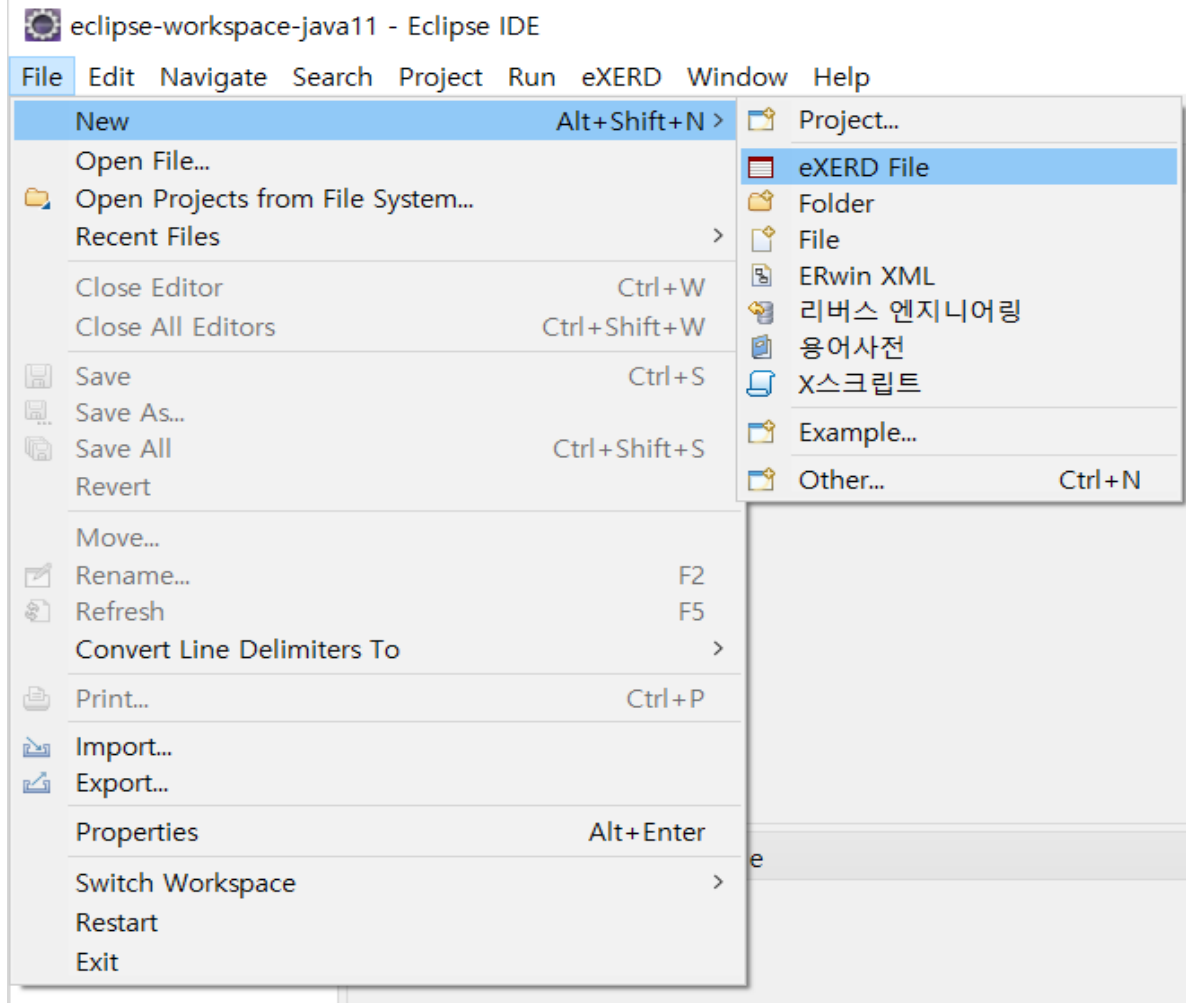


# exERD

## ❖ 이클립스에 exERD 사용하기



## ❖ 이클립스에 exERD 사용하기



## ❖ 이클립스에 exERD 사용하기



### eXERD 새 모델링 파일 위저드

새 모델링 파일을 생성합니다.

대상 DBMS : Oracle 9i ~ 12c

프로젝트 선택

☐ 기존 프로젝트:

- > Servers
- > javaproject
- > javatest

☒ 새로운 프로젝트:

newProject

도메인

☒ 기본 도메인 사용

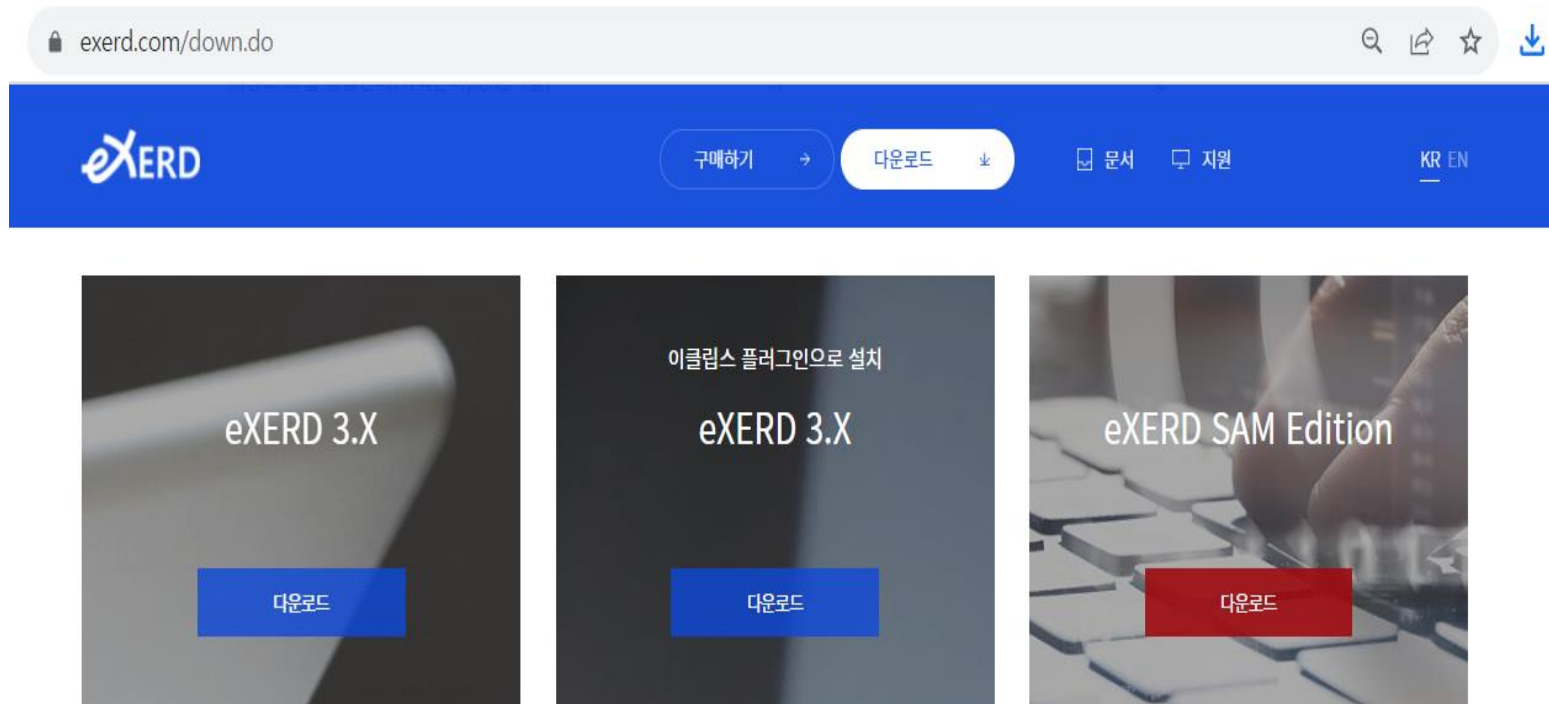
☐ 사용자 지정 도메인 사용

파일 이름: test.exerd

Finish Cancel

## ❖ exERD 다운로드

<https://www.exerd.com/down.do>

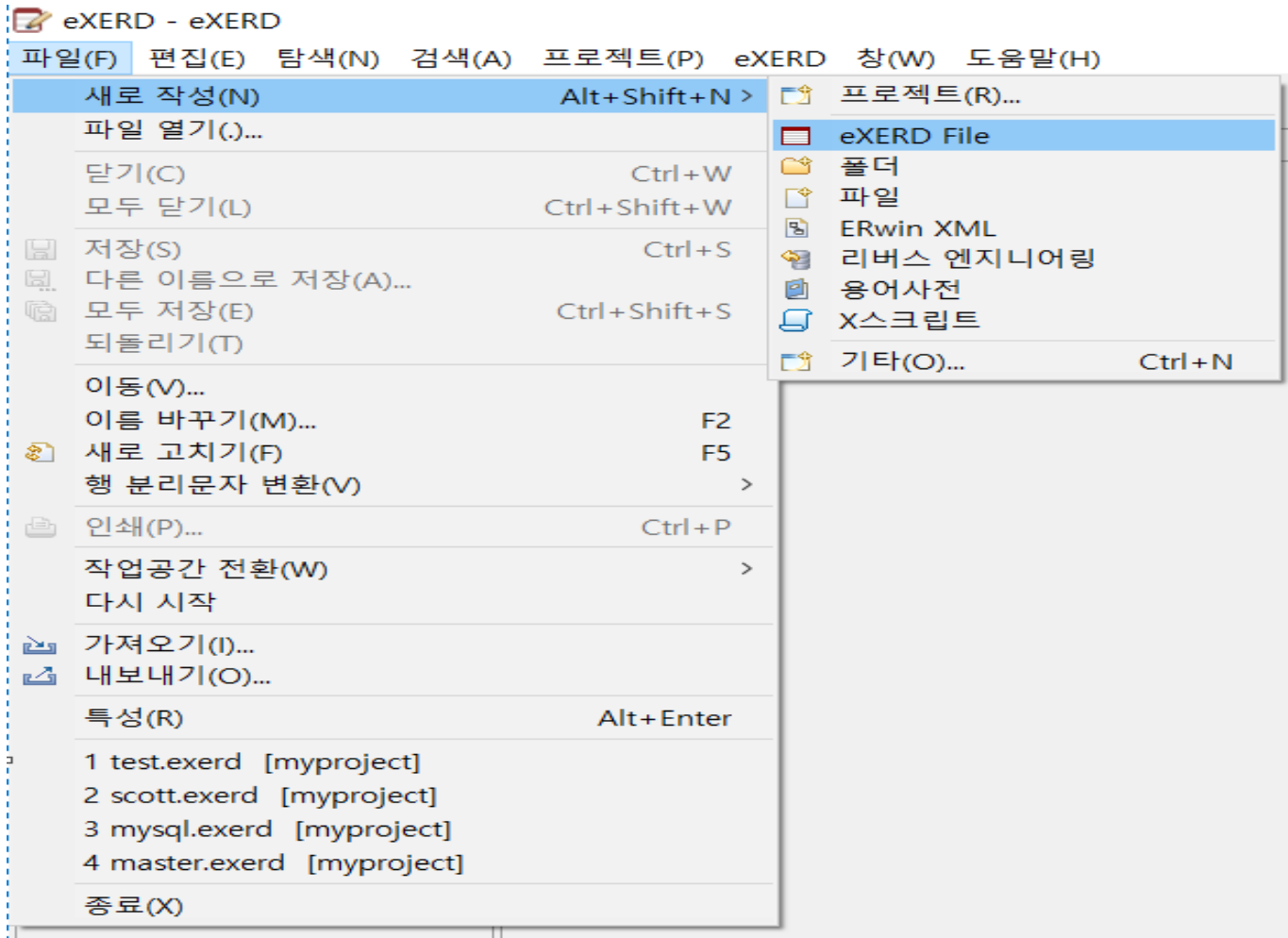


eXERD v2.X 평가판 다운로드




# exERD

## ❖ exERD 실행



## ❖ exERD 실행

 eXERD

eXERD 새 모델링 파일 위저드

새 모델링 파일을 생성합니다.

대상 DBMS : Oracle 9i ~ 12c

프로젝트 선택

☐ 기존 프로젝트:

myproject

☒ 새로운 프로젝트:

newProject

도메인

☒ 기본 도메인 사용

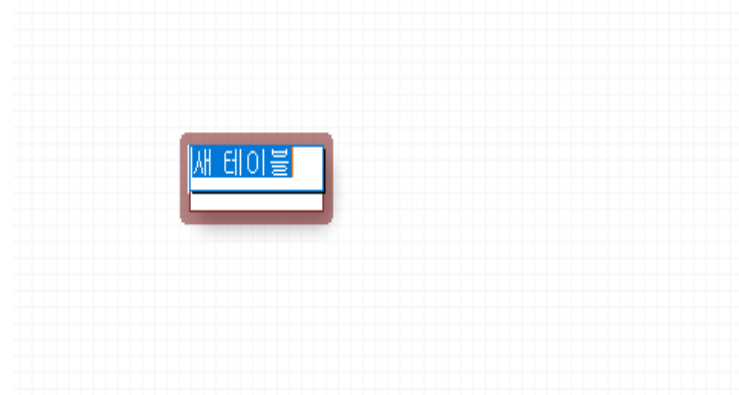
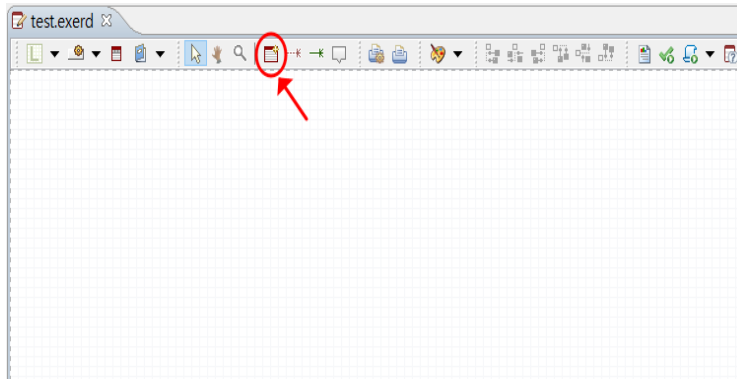
☐ 사용자 지정 도메인 사용

파일 이름: test.exerd

완료(F) 취소

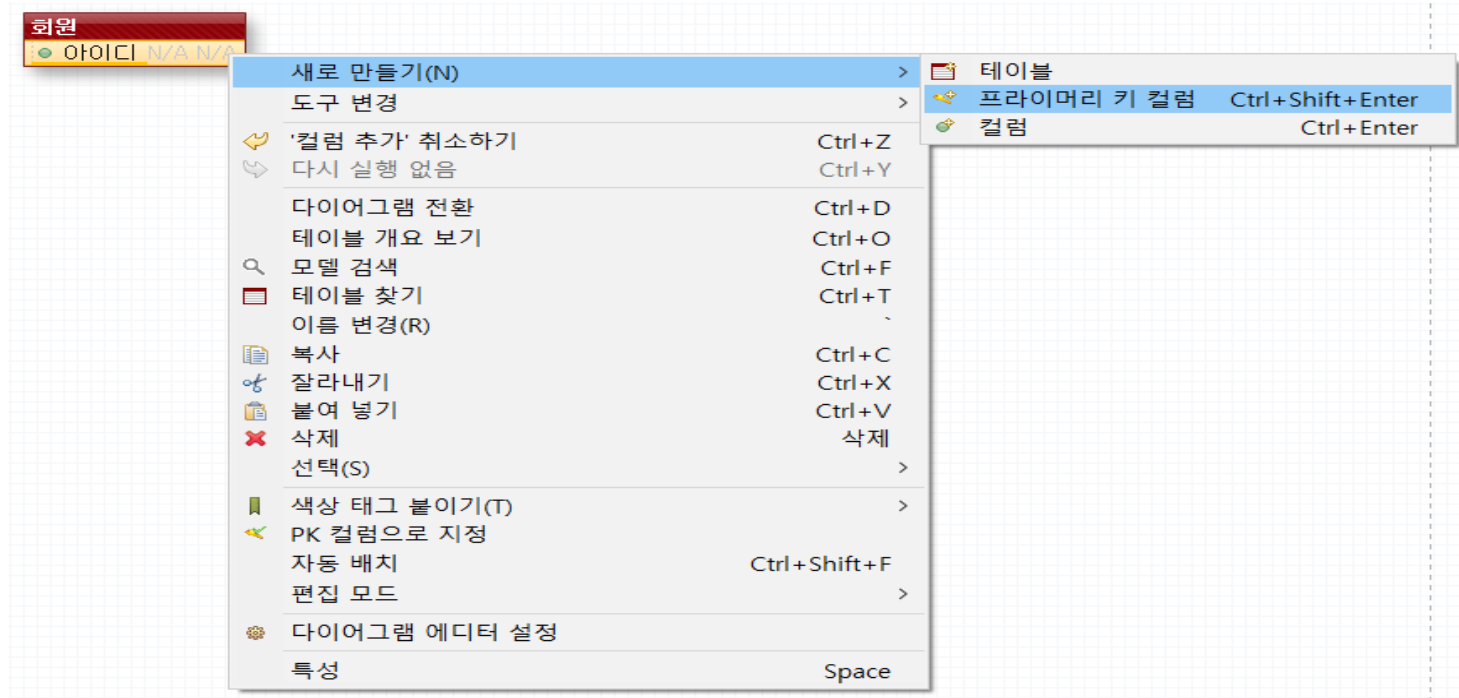
## ❖ 테이블 만들기

- 툴바에서 버튼을 선택하거나 단축키 "3"을 누릅니다.
- 마우스 포인터 모양이 변경되어 표시됩니다.
- 다이어그램의 원하는 위치에서 클릭합니다.



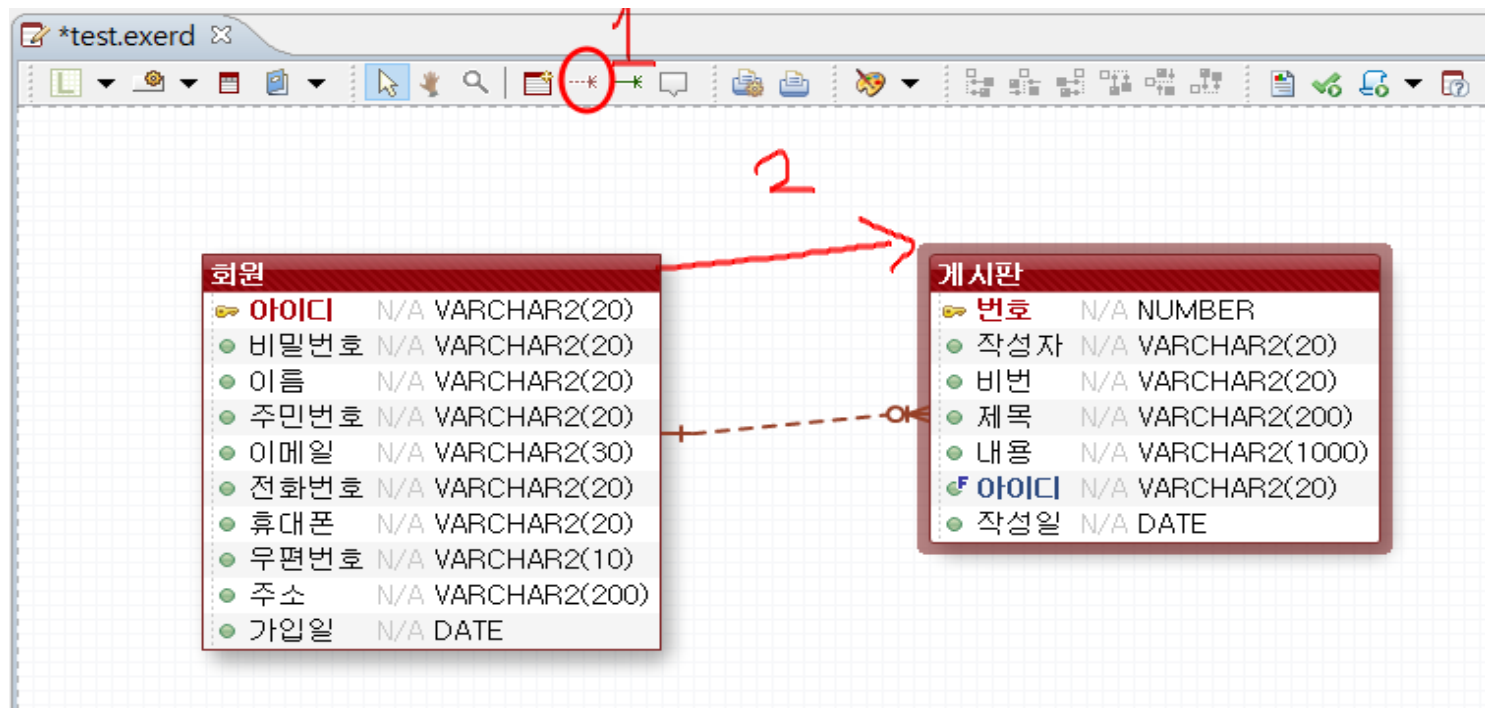
## ❖ 컬럼 추가

- 기본키 추가 ( Ctrl + Shift + Enter )
- 일반 컬럼 추가 ( Ctrl + Enter )
- 팝업메뉴로 추가



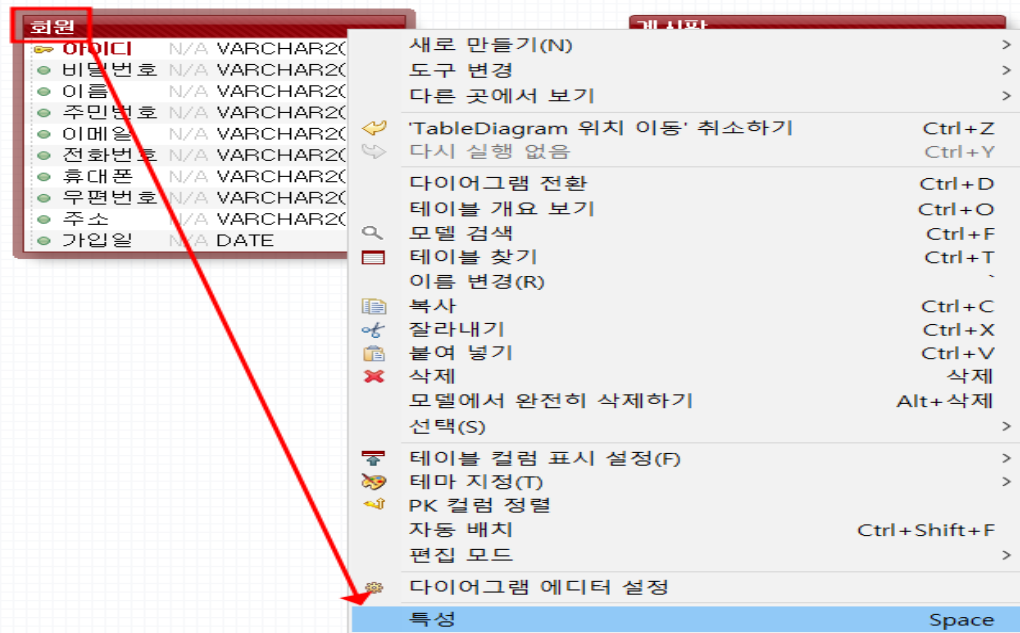
## ❖ 관계설정

- 도구에서 식별관계나 비식별관계를 클릭하고 부모 테이블을 선택하고 자식 테이블로 드래그 하면 됩니다.



## ❖ 물리적 모델링으로 변경

- 테이블을 선택 후 Space 키 또는 팝업메뉴 > 속성을 실행합니다.
- 해당 테이블의 일반적인 속성들과 DBMS의 물리적 속성들을 편집할 수 있습니다.
- 컬럼 및 제약사항, 관계, 인덱스 등 해당 테이블의 모든 내용을 편집할 수 있습니다.



## ❖ 물리적 모델링으로 변경 : 회원(member)

속성

회원

기본 키 제약사항 논리이름 변경 작업이 수행 되었습니다.

일반

컬럼

관계 (외래 키)

제약 사항

인덱스

트리거

기본 키 제약사항

논리 이름

회원

물리 이름

PK\_member

논리 이름	물리 이름	도메인	데이터 타입	NOT NU...	기본 키	기본 값 표...
아이디	id	(None)	VARCHAR2(20)	N/A	PK	(None)
비밀번호	passwd	(None)	VARCHAR2(50)	NULL		(None)
이름	name	(None)	VARCHAR2(50)	NULL		(None)
주민번호	jumin	(None)	VARCHAR2(15)	NULL		(None)
이메일	email	(None)	VARCHAR2(100)	NULL		(None)
전화번호	tel	(None)	VARCHAR2(20)	NULL		(None)
휴대폰	phone	(None)	VARCHAR2(20)	NULL		(None)
우편번호	postcode	(None)	VARCHAR2(10)	NULL		(None)
주소	address	(None)	VARCHAR2(255)	NULL		(None)
가입일	register	(None)	DATE	NULL		(None)

일반

사용자 정의 속성

주석

Default Value Expression :

실행 취소

다시 실행

확인

취소

## ❖ 물리적 모델링으로 변경 : 게시판(board)



게시판

기본 키 제약사항 논리이름 변경 작업이 수행 되었습니다.

일반

컬럼

관계 (외래 키)

제약 사항

인덱스

트리거

기본 키 제약사항

논리 이름

게시판 기본키

물리 이름

PK\_board

논리 이름	물리 이름	도메인	데이터 타입	NOT NU...	기본 키	기본 값 표...
번호	no	(None)	NUMBER	N/A	PK	(None)
작성자	writer	(None)	VARCHAR2(...)	NULL		(None)
비번	password	(None)	VARCHAR2(...)	NULL		(None)
제목	subject	(None)	VARCHAR2(...)	NULL		(None)
내용	content	(None)	VARCHAR2(...)	NULL		(None)
아이디	id	(None)	VARCHAR2(...)	NULL		(None)
작성일	reg_date	(None)	DATE	NULL		(None)

일반

사용자 정의 속성

주석

Default Value Expression :

실행 취소

다시 실행

확인

취소



# 오라클 계정 등록

## ❖ 계정 생성

### 1. 오라클 데이터베이스에 새로운 계정 생성

```
sqlplus system/oracle
```

```
create user master identified by 1234;
```

### 2. 권한 부여

```
grant connect, resource to master;
```

# MySQL 계정 등록

## ❖ 계정 등록

DB명: mydb

user : master

pass : 1234

1. root 계정으로 접속

```
c:\W> mysql -uroot -p1234 mysql
```

2. 새로운 데이터베이스 생성 (mydb)

방법1. console 상태

```
c:\W> mysqladmin -u root -p create mydb
```

방법2. 데이터베이스에 접속된 상태

```
mysql> create database mydb;
```

# MySQL 계정 등록

## 3. 계정 생성 및 권한 부여

### ➤ MySQL 5.7 까지

```
mysql> grant all privileges on mydb.* to master@'%' identified by '1234'  
with grant option;  
mysql> flush privileges;
```

### ➤ MySQL 8 에서는 계정 생성과 DB권한 부여를 각각 수행해야 한다.

```
mysql> create user master@'%' identified by '1234' ;  
mysql> grant all privileges on mydb.* to master@'%' with grant option ;  
mysql> flush privileges ;
```

## 4. 종료

```
mysql> quit;
```

# 포워드 엔지니어링

## ❖ 포워드 엔지니어링 실행 [Exerd] - [포워드 엔지니어링]

포워드 엔지니어링

DDL 생성 옵션

Data Definition Language의 생성 옵션을 지정 해 주세요

설정: 마지막으로 설정 된 구성

공통	
<input checked="" type="checkbox"/> 이름 앞에 스키마 표시	<input type="checkbox"/> 이름을 따옴표(")로
데이터베이스	
<input type="checkbox"/> 테이블 스페이스 생성	<input type="checkbox"/> 테이블 스페이스 식
시퀀스	
<input type="checkbox"/> 시퀀스 생성	<input type="checkbox"/> 시퀀스 삭제
평션 / 프로시저	
<input checked="" type="checkbox"/> 평션 생성	<input type="checkbox"/> 평션 삭제
<input checked="" type="checkbox"/> 프로시저 생성	<input type="checkbox"/> 프로시저 삭제
트리거	
<input checked="" type="checkbox"/> 트리거 생성	<input type="checkbox"/> 트리거 삭제
뷰	
<input checked="" type="checkbox"/> 뷰 생성	<input type="checkbox"/> 뷰 삭제
테이블	
<input checked="" type="checkbox"/> 테이블 생성	<input type="checkbox"/> 테이블 삭제
<input type="checkbox"/> 물리적 특성들 (Storage 및 추가 속성)	

미리 보기:

```
1 /* 내 테이블1 */
2 CREATE TABLE my_table1 (
3     my_pk1 INTEGER NOT NULL, /* 나
4     my_pk2 INTEGER NOT NULL, /* 나
5     my_column1 VARCHAR(20) NOT NU
6     my_column2 VARCHAR(20) /* 내 컬
7 );
8
9 COMMENT ON TABLE my_table1 IS 'DD
10
11 COMMENT ON COLUMN my_table1.my_pk
12
13 COMMENT ON COLUMN my_table1.my_pk
14
15 COMMENT ON COLUMN my_table1.my_co
16
17 COMMENT ON COLUMN my_table1.my_co
18
19 CREATE UNIQUE INDEX my_table1_pk
20 ON my_table1 (
21     my_pk1 ASC,
```

< 이전(B) 다음(N) > 완료(F) 취소

# 포워드 엔지니어링

## ❖ 계정 연결

포워드 엔지니어링

연결 설정

❌ 연결 테스트가 진행되지 않았습니다.

연결:	빈 설정	⚙️ 설정 관리
JDBC 드라이버:		📁 파일 선택
드라이버 클래스:	사용 가능한 드라이버 없음	
URL:	jdbc:oracle:thin:@<host>:1521:<SID> <input type="checkbox"/> 직접 입력	
호스트:		포트: 1521 SID:
사용자:		
비밀번호:		
Connect as	<input checked="" type="radio"/> NORMAL <input type="radio"/> SYSDBA <input type="radio"/> SYSOPER	
연결 테스트		
?	< 이전(B)	다음(N) > 완료(F) 취소

## 계정 연결



# 포워드 엔지니어링

## ❖ 계정 연결

연결 설정

새 드라이버 연결을 설정합니다.

IBM DB2 UDB 8 ~ 11 for  
MySQL 5.1 ~ 5.8  
**Oracle 9i ~ 12c**  
PostgreSQL 9.0 ~ 10.2  
Microsoft SQL Server 200

이름: master

JDBC 드라이버: C:\WoraclexeWappWoracleWproductW 파일 선택

드라이버 클래스: oracle.jdbc.OracleDriver

URL: jdbc:oracle:thin:@localhost:1521:xe ☐ 직접 입력

호스트: localhost 포트: 1521 SID: xe

사용자: master

비밀번호: ●●●●

Connect as ☒ NORMAL ☐ SYSDBA ☐ SYSOPER

**연결 테스트**

확인 취소

# 포워드 엔지니어링

## ❖ 계정 연결

포워드 엔지니어링

연결 설정

① 완료 (FINISH) 버튼을 누르면 DDL 생성을 시작합니다.

연결: master 설정 관리

JDBC 드라이버: C:\Woraclexe\app\Woracle\product\11.2.0\server\jdbc\lib\ojdbc6.jar 파일 선택

드라이버 클래스: oracle.jdbc.OracleDriver

URL: jdbc:oracle:thin:@localhost:1521:xe 직접 입력

호스트: localhost 포트: 1521 SID: xe

사용자: master

비밀번호: ●●●●

Connect as ☒ NORMAL ☐ SYSDBA ☐ SYSOPER

**연결 테스트** →

연결 테스트

연결 성공

확인

② < 이전(B) 다음(N) > **완료(F)** 취소



# 리버스 엔지니어링

- ❖ 리버스 엔지니어링 실행  
[Exerd] – [리버스 엔지니어링]

리버스 엔지니어링

파일 생성

파일 이름:

프로젝트 선택

☒ 기존 프로젝트:

- ☐ myproject
- ☒ newProject

☐ 새로운 프로젝트:

# 리버스 엔지니어링

## ❖ 리버스 엔지니어링 실행 - scott 계정 연결

리버스 엔지니어링

타겟 DBMS에 대한 접속정보를 설정합니다.

① 완료 (FINISH) 버튼을 누르면 리버스 엔지니어링을 시작합니다.

대상 DBMS : Oracle 9i ~ 12c

연결: scott 설정 관리

JDBC 드라이버: C:\Woraclexe\app\oracle\product\11.2.0\se 파일 선택

드라이버 클래스: oracle.jdbc.OracleDriver

URL: jdbc:oracle:thin:@localhost:1521:xe ☐ 직접 입력

호스트: localhost 포트: 1521 SID: xe

사용자: scott

비밀번호: ●●●●●●

Connect as ☒ NORMAL ☐ SYSDBA ☐ SYSOPER

**연결 테스트**

? < 이전(B) **다음(N) >** 완료(F) 취소

# 리버스 엔지니어링

## ❖ 리버스 엔지니어링 실행

리버스 엔지니어링

역공학 옵션 설정

항목 설정	소유자 설정
<input checked="" type="checkbox"/> Tablespace	<input type="radio"/> 접근 가능한 모든 객체
<input checked="" type="checkbox"/> Table	<input checked="" type="radio"/> 현재 접속자가 소유한 객체
<input checked="" type="checkbox"/> View	<input type="radio"/> 스키마 목록에서 선택
<input checked="" type="checkbox"/> Trigger	<input type="text"/> <input type="button" value="선택"/>
<input checked="" type="checkbox"/> Function	
<input checked="" type="checkbox"/> Stored Procedure	
	논리 설정
	<input type="radio"/> eXERD 메타 데이터 사용
	<input checked="" type="radio"/> eXERD 메타 데이터 사용하지 않음
	<input type="radio"/> 물리 이름을 논리 이름으로 사용
	<input checked="" type="radio"/> 주석을 논리 이름으로 사용

# 정규화



# 정규화

## ❖ 정규화(Normalization)

테이블의 중복제거, 데이터 분류 등을 통해 데이터를 구조화 시키는 과정이다.

## ❖ 정규화의 목적

- 중복을 제거하여 공간 절약
- 데이터 무결성을 보장하여 정확성과 일관성을 유지
- 삽입, 삭제, 갱신 이상의 발생 가능성을 줄이는 것이다.

# 정규화

## ❖ 정규화의 장점

1. 데이터베이스 변경 시 이상 현상(Anomaly)을 제거할 수 있다.
2. 정규화된 데이터베이스 구조에서는 새로운 데이터 형의 추가로 인한 확장 시, 그 구조를 변경하지 않아도 되거나 일부만 변경해도 된다.
3. 데이터베이스와 연동된 응용 프로그램에 최소한의 영향만을 미치게 되어 응용프로그램의 생명을 연장시킨다.

## ❖ 정규화의 단점

1. 릴레이션의 분해로 인해 릴레이션 간의 JOIN 연산이 많아진다.
2. 질의에 대한 응답 시간이 느려질 수도 있다. 데이터의 중복 속성을 제거하고 결정자에 의해 동일한 의미의 일반 속성이 하나의 테이블로 집약되므로 한 테이블의 데이터 용량이 최소화되는 효과가 있다.
3. 따라서 데이터를 처리할 때 속도가 빨라질 수도 있고 느려질 수도 있다.

# 정규화

## ❖ 정규화의 종류

제1정규화 ~ 제5정규화 (실무에서는 제3정규화, BCNF까지만 수행)



# 정규화

## ❖ 제1정규화

제1 정규화란 테이블의 컬럼이 원자값(Atomic Value, 하나의 값)을 갖도록 테이블을 분해하는 것이다.

고객취미들(이름, 취미들)

이름	취미들
김연아	인터넷
추신수	영화, 음악
박세리	음악, 쇼핑
장미란	음악
박지성	게임

위의 테이블에서 추신수와 박세리는 여러 개의 취미를 가지고 있기 때문에 제1 정규형을 만족하지 못하고 있다.



# 정규화

## ❖ 제1정규화

제1 정규화를 진행한 테이블은 아래와 같다.

고객취미(이름, 취미)

이름	취미
김연아	인터넷
추신수	영화
추신수	음악
박세리	음악
박세리	쇼핑
장미란	음악
박지성	게임

# 정규화

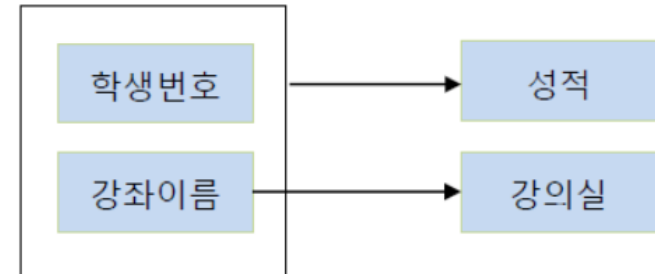
## ❖ 제2정규화

제2 정규화란 제1 정규화를 진행한 테이블에 대해 완전 함수 종속을 만족하도록 테이블을 분해하는 것이다.

여기서 완전 함수 종속이라는 것은 기본키의 부분집합이 결정자가 되어서는 안된다는 것을 의미한다.

수강강좌

학생번호	강좌이름	강의실	성적
501	데이터베이스	공학관 110	3.5
401	데이터베이스	공학관 110	4.0
402	스포츠경영학	체육관 103	3.5
502	자료구조	공학관 111	4.0
501	자료구조	공학관 111	3.5



1. 이 테이블에서 기본키는 (학생번호, 강좌이름)으로 복합키이다.
2. (학생번호, 강좌이름)인 기본키는 성적을 결정하고 있다.
3. (학생번호, 강좌이름) --> (성적) 그런데 여기서 강의실이라는 컬럼은 기본키의 부분 집합인 강좌이름에 의해 결정될 수 있다.

# 정규화

## ❖ 제2정규화

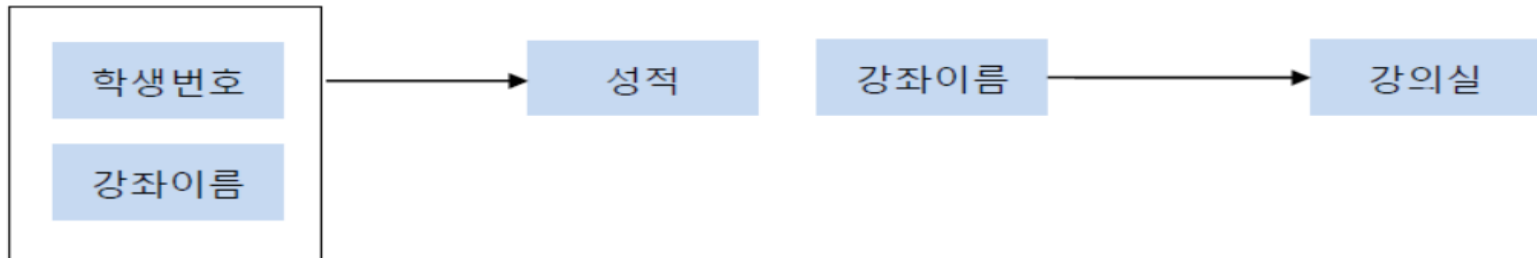
(강좌이름) --> (강의실) 즉, **기본키(학생번호, 강좌이름)의 부분키인 강좌이름이 결정자**이기 때문에 위의 테이블의 경우 다음과 같이 기존의 테이블에서 강의실을 분해하여 별도의 테이블로 관리하여 제2 정규형을 만족시킬 수 있다.

수강

학생번호	강좌이름	성적
501	데이터베이스	3.5
401	데이터베이스	4.0
402	스포츠경영학	3.5
502	자료구조	4.0
501	자료구조	3.5

강의실

강좌이름	강의실
데이터베이스	공학관 110
스포츠경영학	체육관 103
자료구조	공학관 111



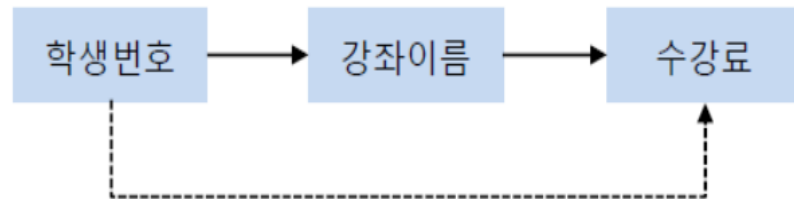
# 정규화

## ❖ 제3정규화

제3 정규화란 제2 정규화를 진행한 테이블에 대해 이행적 종속을 없애도록 테이블을 분해하는 것이다. 여기서 이행적 종속이라는 것은  $A \rightarrow B$ ,  $B \rightarrow C$ 가 성립할 때  $A \rightarrow C$ 가 성립되는 것을 의미한다.

계절학기

학생번호	강좌이름	수강료
501	데이터베이스	20000
401	데이터베이스	20000
402	스포츠경영학	15000
502	자료구조	25000



기존의 테이블에서 학생 번호는 강좌 이름을 결정하고 있고, 강좌 이름은 수강료를 결정하고 있다. 그렇기 때문에 (학생 번호, 강좌 이름) 테이블과 (강좌 이름, 수강료) 테이블로 분해해야 한다.

# 정규화

## ❖ 제3정규화

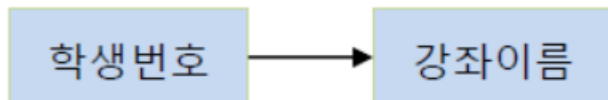
학생 번호를 통해 강좌 이름을 참조하고, 강좌 이름으로 수강료를 참조하도록 테이블을 분해해야 하며 이행적 종속을 없앤 결과는 다음의 그림과 같다.

계절수강

학생번호	강좌이름
501	데이터베이스
401	데이터베이스
402	스포츠경영학
502	자료구조

수강료

강좌이름	수강료
데이터베이스	20000
스포츠경영학	15000
자료구조	25000



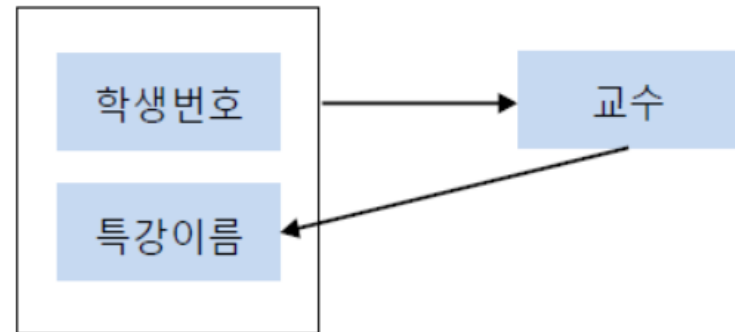
# 정규화

## ❖ BCNF 정규화

BCNF 정규화란 제3 정규화를 진행한 테이블에 대해 모든 결정자가 후보키가 되도록 테이블을 분해하는 것이다.

특강수강

학생번호	특강이름	교수
501	소셜네트워크	김교수
401	소셜네트워크	김교수
402	인간과 동물	승교수
502	창업전략	박교수
501	창업전략	홍교수



- 1.특강수강 테이블에서 기본키는 (학생번호, 특강이름)이다.
- 2.기본키 (학생번호, 특강이름)는 교수를 결정하고 있다. 또한 여기서 교수는 특강이름을 결정하고 있다.
- 3.그런데 문제는 **교수가 특강이름을 결정하는 결정자이지만, 후보키가 아니라는 점**이다.
- 4.그렇기 때문에 BCNF 정규화를 만족시키기 위해서 위의 테이블을 분해해야 하는데, 특강신청 테이블과 특강교수 테이블로 분해할 수 있다.

# 정규화

## ❖ BCNF 정규화

BCNF 정규화를 만족시키기 위해서 위의 테이블을 분해해야 하는데, 다음과 같이 특강신청 테이블과 특강교수 테이블로 분해할 수 있다.

특강신청

학생번호	교수
501	김교수
401	김교수
402	승교수
502	박교수
501	홍교수

특강교수

특강이름	교수
소셜네트워크	김교수
인간과 동물	승교수
창업전략	박교수
창업전략	홍교수

학생번호

교수

특강이름

교수