



빅데이터 모델링과 플랫폼 = 빅모플

# Big Data Modeling and Platforms

AIX30011 Spring 2025

**한동대학교**  
HANDONG GLOBAL UNIVERSITY

School of Applied Artificial Intelligence | 전재영



# Copyright Notice

These slides are available for educational purposes. You may not use or distribute these slides for commercial purposes. You may make copies of these slides and use or distribute them for educational purposes as long as you cite the author as the source of the slides.

# 게시판

The screenshot shows a web browser window with the URL `lms.handong.edu/courses/17877/external_tools/39`. The page is titled 'Assignment 01 - MySQL, DataGrip, AirportDB'. It features a sidebar with navigation links like '홈', '수업 계획서', '공지', '주차학습', '출결현황', '과제 및 평가', '시험 및 설문', '게시판', '파일', 'Zoom', '토론', '사용자 및 그룹', and '성적'. The main content area displays a list of forum posts with columns for 'No', '제목', '북마크', '작성자', '등록일시', '조회수', and '통계'. The posts are numbered 1 to 5, with the most recent at the top. Each post has a '보기' button next to it. The interface is in Korean and includes various UI elements like search bars, filters, and pagination controls.

2025년 1학기

≡ 전체 게시판 목록 바로가기    다른 게시판으로 이동 ▾

### Assignment 01 - MySQL, DataGrip, AirportDB

게시물 모두 가져오기    모든 게시물 첨부파일 다운로드    설정    ✎ 글쓰기

검색결과 5 개    ☐ 안 읽은 게시물    ☐ 북마크 게시물    제목 ▾    검색어 입력    🔍


No	제목	북마크	작성자	등록일시	조회수	통계
5	최종 확인은... (0) <span>N</span>		전재영	2025-03-27 15:26	23	<a href="#">보기</a>
4	MySQL: Access denied for user 'root' (0) <span>N</span>		전재영	2025-03-27 15:23	23	<a href="#">보기</a>
3	No homebrew (1) <span>N</span>		전재영	2025-03-27 15:17	27	<a href="#">보기</a>
2	Host / Guest 구분 필요 (0) <span>N</span>		전재영	2025-03-27 11:23	29	<a href="#">보기</a>
1	9시간 🍷 (0) <span>N</span>		전재영	2025-03-27 11:16	25	<a href="#">보기</a>

< 1 >

# Office Hours

수업 관련 질문 없더라도, 학기초 반드시 면담 필요

<https://calendly.com/hisplan-handong/casual-chat>



Jaeyoung Chun (전재영)

## Office Hours

🕒 30 min

📍 Ebenezer 310

### Select a Date & Time

< April 2025 > Thursday, April 10

MON	TUE	WED	THU	FRI	SAT	SUN
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

11:30

15:30

16:30

### Time zone

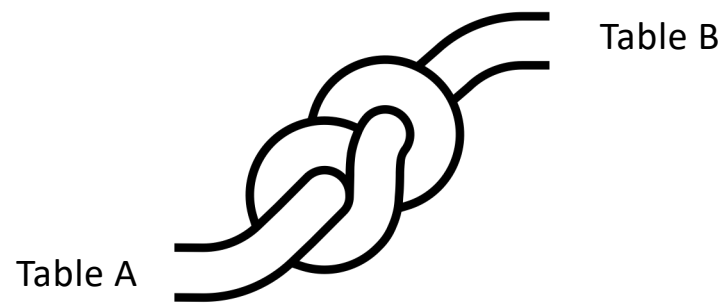
🌐 Japan, Korea Time (21:22) ▼

Jaeyoung Chun

# SQL

1. INNER JOIN
2. LEFT OUTER JOIN
3. INSERT
4. UPDATE
5. DELETE
6. TRUNCATE TABLE

7. CREAT DATABASE
8. SHOW DATABASES
9. USE
10. SHOW TABLES
11. CREATE TABLE
12. DROP TABLE
13. ALTER TABLE (ADD PRIMARY KEY)
14. ALTER TABLE (ADD COLUMN)
15. ALTER TABLE (DROP COLUMN)
16. ALTER TABLE (ADD UNIQUE)
17. ALTER TABLE (DROP INDEX)



# SQL Join

# INNER JOIN

Left Table

Date	CountryID	Units
1/1/2020	1	40
1/2/2020	1	25
1/3/2020	3	30
1/4/2020	2	35

Right Table

ID	Country
3	Panama
4	Spain

Merged Table

Date	CountryID	Units	Country
1/3/2020	3	30	Panama

```
SELECT
    A.Date, A.CountryID, A.Units,
    B.Country
FROM left_table AS A
INNER JOIN right_table AS B
ON A.CountryID = B.ID
;
```

(\*) 컬럼 이름이 같을 필요는 없음

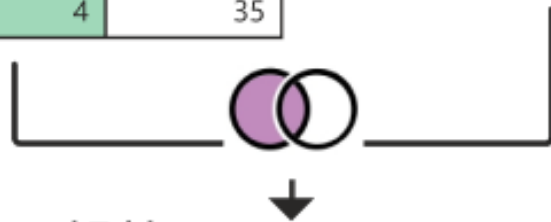
# LEFT OUTER JOIN

Left Table

Date	CountryID	Units
1/1/2020	1	40
1/2/2020	1	25
1/3/2020	3	30
1/4/2020	4	35

Right Table

ID	Country
1	USA
2	Canada
3	Panama



Merged Table

Date	CountryID	Units	Country
1/1/2020	1	40	USA
1/2/2020	1	25	USA
1/3/2020	3	30	Panama
1/4/2020	4	35	<i>null</i>

```
SELECT
    A.Date, A.CountryID, A.Units,
    B.Country
FROM left_table AS A
LEFT OUTER JOIN right_table AS B
ON A.CountryID = B.ID
;
```



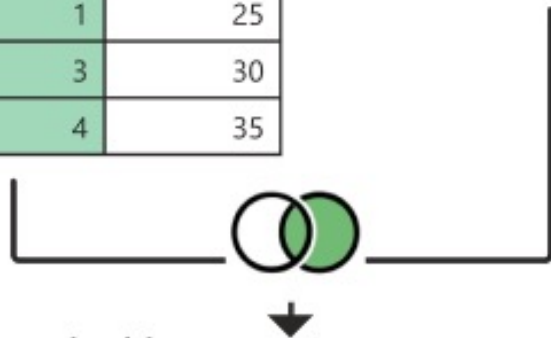
# RIGHT OUTER JOIN

Left Table

Date	CountryID	Units
1/1/2020	1	40
1/2/2020	1	25
1/3/2020	3	30
1/4/2020	4	35

Right Table

ID	Country
3	Panama



Merged Table

Date	CountryID	Units	Country
1/3/2020	3	30	Panama

```
SELECT
    A.Date, A.CountryID, A.Units,
    B.Country
FROM left_table AS A
RIGHT OUTER JOIN right_table AS B
ON A.CountryID = B.ID
;
```

# INNER JOIN

Live



1. 주문이 보류된 경우, 고객에게 전화

orders	orderNumber	orderDate	requiredDate	shippedDate	comments	status	customerNumber
	10334	2004-11-19	2004-11-28	<null>	The outstaniding ...	On Hold	144
	10401	2005-04-03	2005-04-14	<null>	Customer credit l...	On Hold	328
	10407	2005-04-22	2005-05-04	<null>	Customer credit l...	On Hold	450
	10414	2005-05-06	2005-05-13	<null>	Customer credit l...	On Hold	362

▼	orders
>	columns 7
>	keys 1
▼	foreign keys 1
	orders_ibfk_1 (customerNumber) → customers (customerNumber)
>	indexes 2

# LEFT OUTER JOIN

customers	
customerNumber	customerName
103	Atelier graphique
112	Signal Gift Stores
114	Australian Collectors, Co.
119	La Rochelle Gifts
121	Baane Mini Imports
124	Mini Gifts Distributors Ltd.
125	Havel & Zbyszek Co
128	Blauer See Auto, Co.
129	Mini Wheels Co.

1. 한 번도 주문을 해 본적 없는 고객은?

Live



orders			
customerNumber	orderNumber	orderDate	status
103	10123	2003-05-20	Shipped
103	10298	2004-09-27	Shipped
103	10345	2004-11-25	Shipped
112	10124	2003-05-21	Shipped
112	10278	2004-08-06	Shipped
112	10346	2004-11-29	Shipped
114	10120	2003-04-29	Shipped
114	10125	2003-05-21	Shipped
114	10223	2004-02-20	Shipped
114	10342	2004-11-24	Shipped
114	10347	2004-11-29	Shipped
119	10275	2004-07-23	Shipped
119	10315	2004-10-29	Shipped
119	10375	2005-02-03	Shipped
119	10425	2005-05-31	In Process
121	10103	2003-01-29	Shipped
121	10158	2003-10-10	Shipped
121	10309	2004-10-15	Shipped
121	10325	2004-11-05	Shipped
124	10113	2003-03-26	Shipped
124	10135	2003-07-02	Shipped
124	10142	2003-08-08	Shipped
124	10182	2003-11-12	Shipped
124	10229	2004-03-11	Shipped
124	10271	2004-07-20	Shipped
124	10282	2004-08-20	Shipped
124	10312	2004-10-21	Shipped
124	10335	2004-11-19	Shipped
124	10357	2004-12-10	Shipped
124	10368	2005-01-19	Shipped
124	10371	2005-01-23	Shipped
124	10382	2005-02-17	Shipped
124	10385	2005-02-28	Shipped
124	10390	2005-03-04	Shipped
124	10396	2005-03-23	Shipped
124	10421	2005-05-29	In Process
128	10101	2003-01-09	Shipped
128	10230	2004-03-15	Shipped
128	10300	2003-10-04	Shipped
128	10323	2004-11-05	Shipped
129	10111	2003-03-25	Shipped

125?

The slide features three decorative curved lines. One is in the top right corner, curving from the top towards the right. Another is in the bottom left corner, curving from the left towards the bottom. The third is on the left side, curving from the middle towards the bottom. All three lines have a gradient from light blue to light green.

# 여러가지 시나리오를 통한 설계

# Modeling

- 제품 한 개를 가지고 시작한 스타트업
- 고객 관리를 위한 고객 데이터베이스 필요
- 비즈니스가 발전/변화하면서 요구사항도 같이 변화

# Rule of Thumb

## 디자인 원칙 / 가이드라인




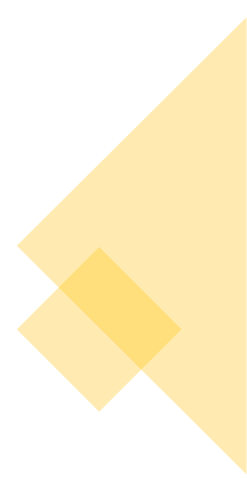
1. 중복 피하기
2. 확장성
3. 한 컬럼에는 한 값만
4. 한 테이블에는 관련 있는 컬럼들만
5. Row는 무제한, Column은 제한
6. 테이블 이름은 복수
7. 컬럼 이름은 단수
8. ...

# CREATE TABLE

테이블 이름  
↓

```
CREATE TABLE customers  
(  
    customerNumber INT NOT NULL,  
    customerName VARCHAR(50) NOT NULL,  
    dateOfBirth DATE NULL  
);
```

↑ 컬럼 이름      ↑ 컬럼 데이터형      ↑ 필수 입력?



# INSERT

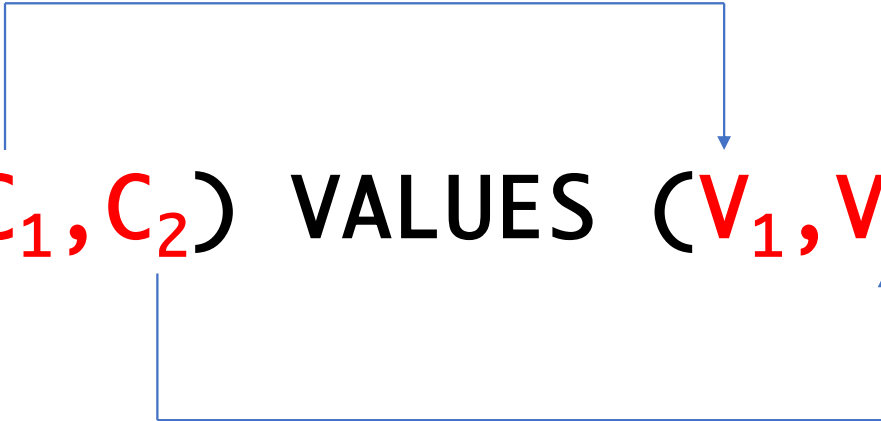




# INSERT

INSERT INTO ... ( $C_1, C_2$ ) VALUES ( $V_1, V_2$ );

짜이 맞아야함



INSERT INTO ... ( $C_1, C_2$ ) VALUES ( $V_1$ ); 🐱

INSERT INTO ... ( $C_1$ ) VALUES ( $V_1, V_2$ ); 🐱

# 처음에는 관참음

Live



customers					
customerNumber	customerName	...	phone	city	...
129	Mini Wheels Co.	...	650222	SF	...
131	Land of Toys	...	212555	NYC	...

# 같은 고객이 계속 구매

Live



customers					
customerNumber	customerName	...	phone	city	...
129	Mini Wheels Co.	...	650222	SF	...
131	Land of Toys	...	212555	NYC	...
129	Mini Wheels Co.	...	650222	SF	...
129	Mini Wheels Co.	...	650222	SF	...

중복 데이터 발생



# UPDATE



# UPDATE

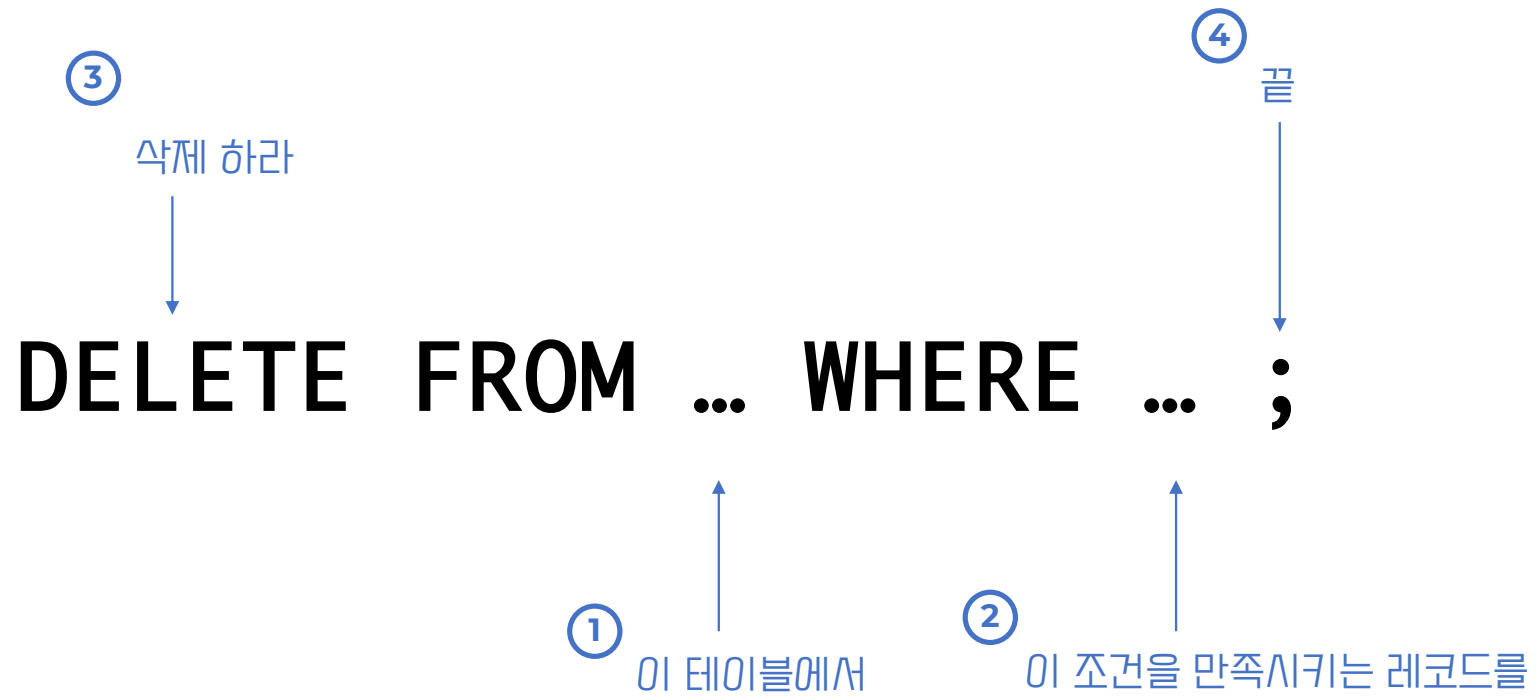
Live



customers					
customerNumber	customerName	...	phone	city	...
129	Mini Wheels Co.	...	650222	SF	...
131	Land of Toys	...	212555	NYC	...
129	Mini Wheels Co.	...	650222	SF	...
129	Mini Wheels Co.	...	650222	SF	...


# DELETE

Live




# 같은 고객이 계속 구매

customers					
customerNumber	customerName	...	phone	city	...
129	Mini Wheels Co.	...	650222	SF	...
129	Mini Wheels Co.	...	650222	SF	...
129	Mini Wheels Co.	...	650222	SF	...



customerNumber	customerName	...	phone	city	...	orderCount
129	Mini Wheels Co.	...	650222	SF	...	3




제품 구매 횟수 추가


# 제품 구매 날짜 추적

봄,여름,가을,겨울 언제 잘 팔리나?

customers					
customerNumber	customerName	...	phone	city	...



customerNumber	customerName	...	phone	city	...	orderDate



제품 구매 날짜 추가



# 총 판매수

```
total_orders = 0
for row in customers:
    dates = row.orderDate.split(",")
    n = len(dates)
    total_orders += n
```

customers						
customerNumber	customerName	...	phone	city	...	orderDate
129	Mini Wheels Co.	...	650222	SF	...	2023-03-24

↓ 추가 구매

customerNumber	customerName	...	phone	city	...	orderDate
129	Mini Wheels Co.	...	650222	SF	...	2023-03-24,2023-04-01,2023-04-15
131	Land of Toys	...	212555	NYC	...	2023-04-01,2023-05-01

SELECT COUNT(\*) FROM customers;



Date?  
Varchar(50)?  
Varchar(100)?

# 4월 이후 구매한 횟수

customers						
customerNumber	customerName	...	phone	city	...	orderDate
129	Mini Wheels Co.	...	650222	SF	...	2023-03-24

↓ 추가 구매

customerNumber	customerName	...	phone	city	...	orderDate
129	Mini Wheels Co.	...	650222	SF	...	2023-03-24, 2023-04-01, 2023-04-15
131	Land of Toys	...	212555	NYC	...	2023-04-01, 2023-05-01

```
SELECT * FROM customers  
WHERE orderDate >= "2023-04-01";
```



# 4월 이후 구매한 횟수

customers						
customerNumber	customerName	...	phone	city	...	orderDate
129	Mini Wheels Co.	...	650222	SF	...	2023-03-24

↓ 추가 구매

컬럼 추가?

customerNumber	customerName	...	phone	city	...	orderDate	orderDate2
129	Mini Wheels Co.	...	650222	SF	...	2023-03-24	2023-04-01

orderDate3, orderDate4, orderDate5, orderDate6, orderDate7, orderDate8, orderDate9, orderDate10,  
orderDate11, orderDate12, orderDate13, ...



```
SELECT * FROM customers
WHERE orderDate >= "2023-04-01" OR orderDate2 >= "2023-04-01" OR ...
```

# 디자인 0

customers						
customerNumber	customerName	...	phone	city	...	orderDate
129	Mini Wheels Co.	...	650222	SF	...	2023-03-24

↓ 추가 구매

customerNumber	customerName	...	phone	city	...	orderDate
129	Mini Wheels Co.	...	650222	SF	...	2023-03-24
129	Mini Wheels Co.	...	650222	SF	...	2023-04-01
129	Mini Wheels Co.	...	650222	SF	...	2023-04-15

중복 데이터 발생



# 디자인 1

Live



customers						
customerNumber	customerName	...	phone	city	...	
129	Mini Wheels Co.	...	650222	SF	...	
131	Land of Toys	...	212555	NYC	...	

orders	
customerNumber	orderDate
129	2023-03-24
129	2023-04-01
129	2023-04-15
131	2023-04-01

Join Query

A	customerNumber	customerName	orderDate
	129	Mini Wheels Co.	2023-03-24
	129	Mini Wheels Co.	2023-04-01
	129	Mini Wheels Co.	2023-04-15
	131	Land of Toys	2023-04-01

B	customerNumber	orderCount
	129	3
	131	1

# 디자인 1

Live



customers						
customerNumber	customerName	...	phone	city	...	
129	Mini Wheels Co.	...	650222	SF	...	
131	Land of Toys	...	212555	NYC	...	

orders	
customerNumber	orderDate
129	2023-03-24
129	2023-04-01
129	2023-04-15
131	2023-04-01
999	2023-04-12

고객번호 999  
존재하지 않아도 입력 가능



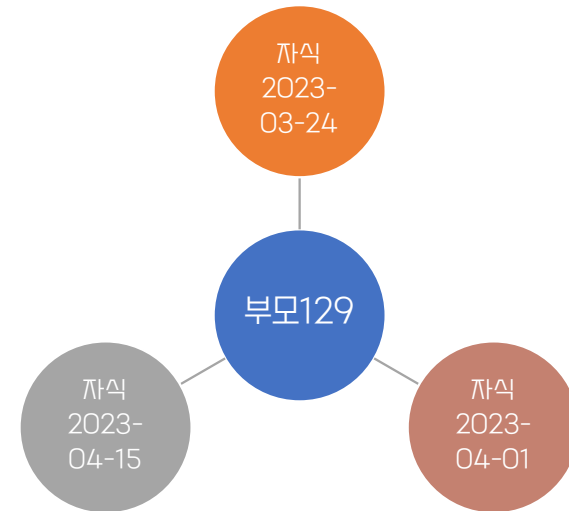
# 부모/자식

부모129

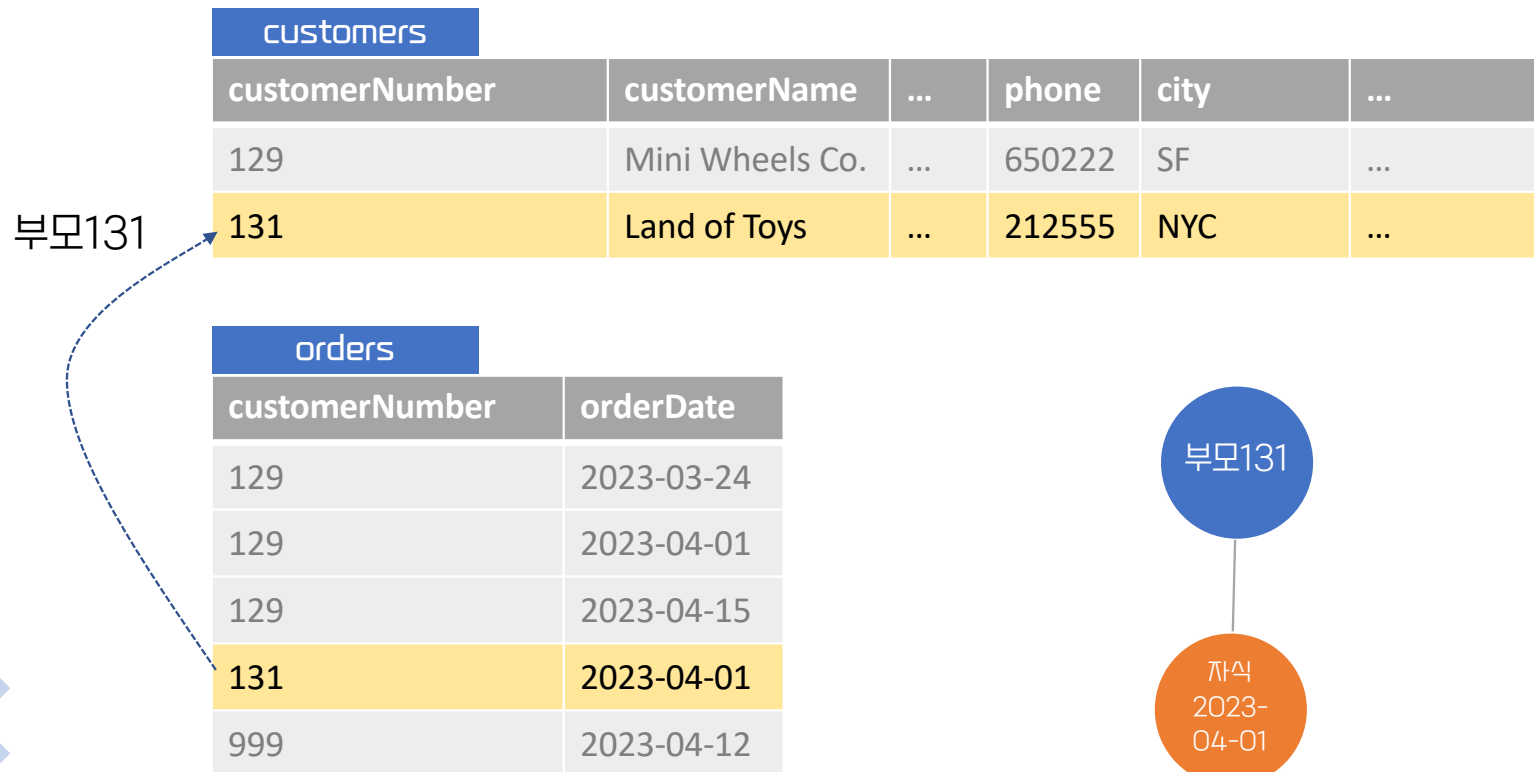
customers						
customerNumber	customerName	...	phone	city	...	
129	Mini Wheels Co.	...	650222	SF	...	
131	Land of Toys	...	212555	NYC	...	

orders	
customerNumber	orderDate
129	2023-03-24
129	2023-04-01
129	2023-04-15
131	2023-04-01
999	2023-04-12



# 부모/자식





# 부모/자식/고아



customers					
customerNumber	customerName	...	phone	city	...
129	Mini Wheels Co.	...	650222	SF	...
131	Land of Toys	...	212555	NYC	...

orders	
customerNumber	orderDate
129	2023-03-24
129	2023-04-01
129	2023-04-15
131	2023-04-01
999	2023-04-12

부모999는 없는데?



orphan record

# 고아 찾기

Live



customers					
customerNumber	customerName	...	phone	city	...
129	Mini Wheels Co.	...	650222	SF	...
131	Land of Toys	...	212555	NYC	...

orders	
customerNumber	orderDate
129	2023-03-24
129	2023-04-01
129	2023-04-15
131	2023-04-01
999	2023-04-12

1. SELECT 문으로 orphan record를 찾는다.
2. DELETE 문으로 삭제한다.

# 디자인 2

Live



customers						
customerNumber	PK	customerName	...	phone	city	...
129		Mini Wheels Co.	...	650222	SF	...
131		Land of Toys	...	212555	NYC	...

orders	
customerNumber	FK
129	2023-03-24
129	2023-04-01
129	2023-04-15
131	2023-04-01

PK (Primary Key)

동일한 고객번호 131로 고객 테이블에 입력 불가능.

FK (Foreign Key)

고객번호 999로 주문입력 불가능.

# 디자인 2

customers						
customerNumber	PK	customerName	...	phone	city	...
129		Mini Wheels Co.	...	650222	SF	...
131		Land of Toys	...	212555	NYC	...

orders		
customerNumber	FK	orderDate
129		2023-03-24
129		2023-04-01
129		2023-04-15
131		2023-04-01
129		2023-04-15

만약 고객번호 129번이 2023년 4월 15일  
주문을 또 한다면?

이 두 주문을 어떻게 구분하나?



# 디자인 3

Live



customers						
customerNumber	PK	customerName	...	phone	city	...
129		Mini Wheels Co.	...	650222	SF	...
131		Land of Toys	...	212555	NYC	...

orders			
orderNumber	PK	customerNumber	FK
100		129	
110		129	
120		129	
130		131	
140		129	

동일한 주문번호로 재입력 불가능

같은 고객, 같은 주문 날짜이지만,  
주문번호 120

같은 고객, 같은 주문 날짜이지만,  
주문번호 140

# 디자인 4

customers						
customerNumber	PK	customerName	...	phone	city	...
129		Mini Wheels Co.	...	650222	SF	...
131		Land of Toys	...	212555	NYC	...

orders					
orderNumber	PK	customerNumber	FK	orderDate	product
100		129		2023-03-24	A
110		129		2023-04-01	A,B,B,C
120		129		2023-04-15	C,C
130		131		2023-04-01	A,F
140		129		2023-04-15	C,F

만약 판매하는 제품이 여러개 생기고,  
한 번 주문할 때 여러 제품을 동시에  
구문한다면?



# 디자인 4

customers						
customerNumber	PK	customerName	...	phone	city	...
129		Mini Wheels Co.	...	650222	SF	...
131		Land of Toys	...	212555	NYC	...

orders			
orderNumber	PK	customerNumber	FK
100		129	
110		129	
120		129	
130		131	
140		129	

orderdetails		
orderNumber	FK	product
100		A
...		...
120		C
140		C
140		F