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

Jaeyoung Chun

Office Hours

수업 관련 질문 없더라도, 학기초 반드시 면담 필요

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)

Jaeyoung Chun

Table B

Table A
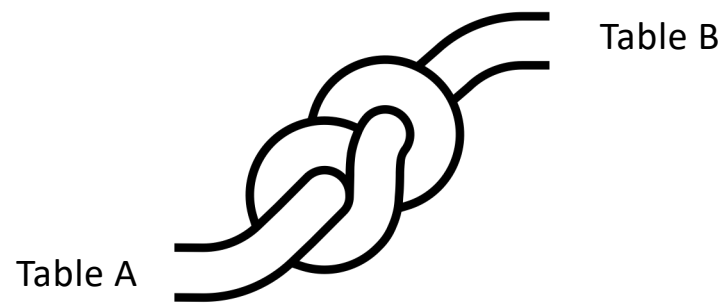
SQL Join

Jaeyoung Chun

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

Jaeyoung Chun

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
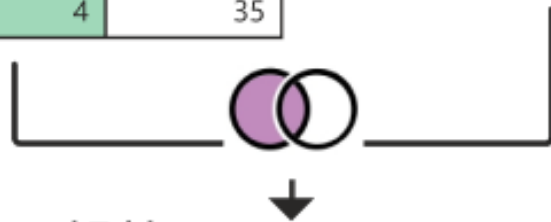
Merged Table

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	*null*

Source: https://learn.microsoft.com/en-us/power-query/

```sql
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
;
```

Jaeyoung Chun

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
;
```

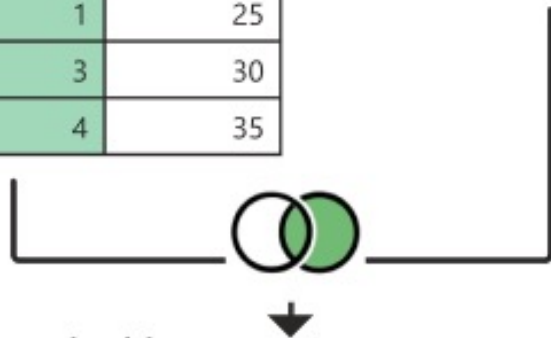Jaeyoung Chun

INNER JOIN

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
```

Jaeyoung Chun

여러가지 시나리오를 통한 설계

Jaeyoung Chun

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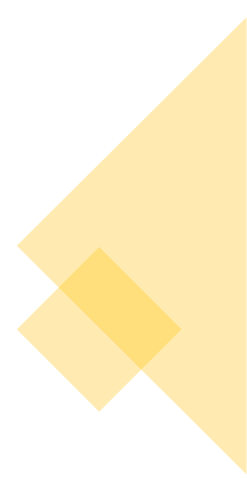
1. 중복 피하기
2. 확장성
3. 한 컬럼에는 한 값만
4. 한 테이블에는 관련 있는 컬럼들만
5. Row는 무제한, Column은 제한
6. 테이블 이름은 복수
7. 컬럼 이름은 단수
8. …

Jaeyoung Chun

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?

Jaeyoung Chun

INSERT

④
저장 하라

③
이 컬럼들에 해당하는

⑤
끝

INSERT INTO … (…) VALUES (…) ;

①
이 테이블에

②
이 값들로
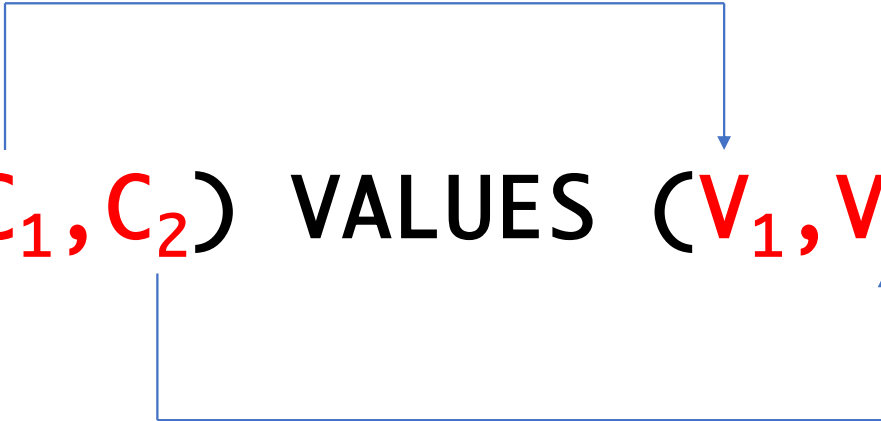
INSERT

짝이 맞아야함

INSERT INTO … (C_1, C_2) VALUES (V_1, V_2);

INSERT INTO … (C_1, C_2) VALUES (V_1); 🙀

INSERT INTO … (C_1) VALUES (V_1, V_2); 🙀

처음에는 괜찮음

customers

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

Jaeyoung Chun

같은 고객이 계속 구매

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

🙀

Jaeyoung Chun

UPDATE

④
업데이트 하라

③
이 내용으로

⑤
끝

UPDATE ... SET ... WHERE ... ;

①
이 테이블에서

②
이 조건을 만족시키면

Jaeyoung Chun

UPDATE

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	...

Jaeyoung Chun

DELETE

③
삭제 하라

④
끝

DELETE FROM … WHERE … ;

①
이 테이블에서

②
이 조건을 만족시키는 레코드를

Jaeyoung Chun
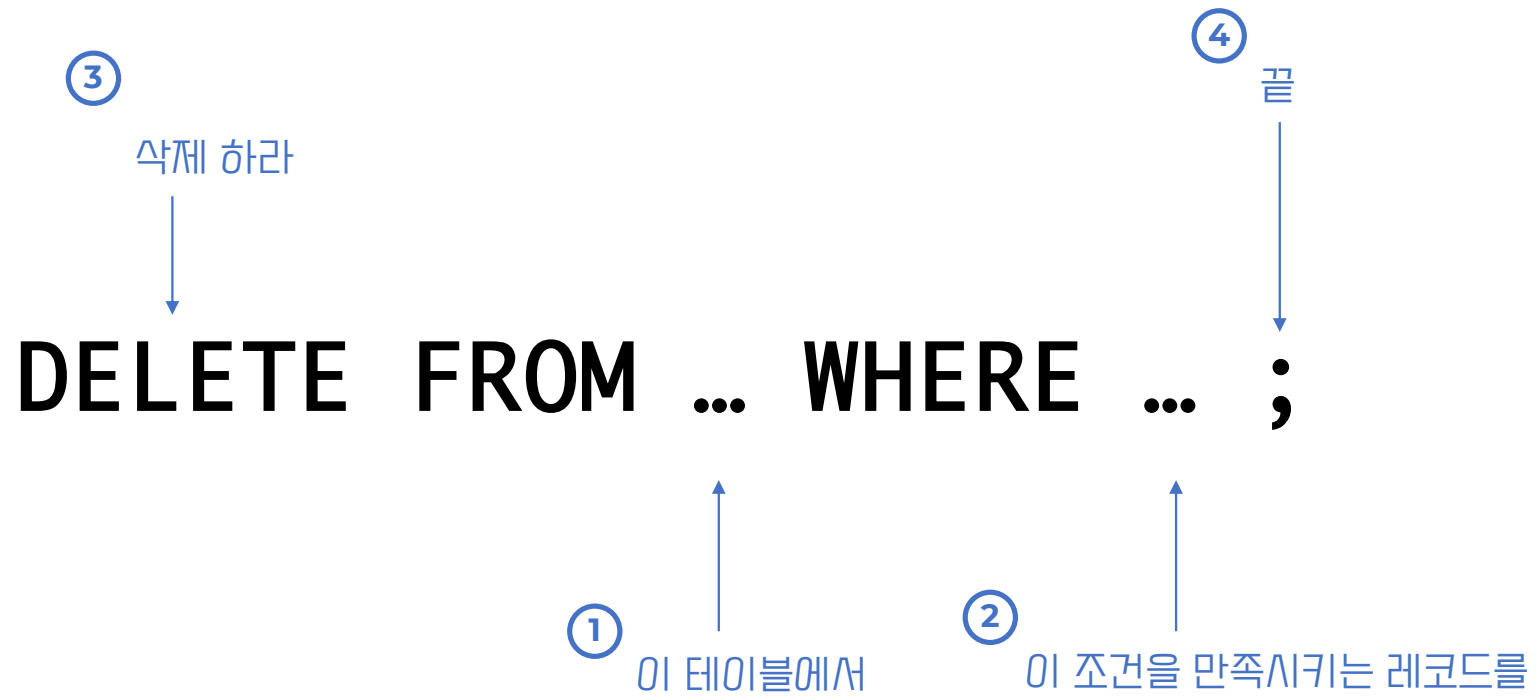
같은 고객이 계속 구매

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

Jaeyoung Chun

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)?

Jaeyoung Chun

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";
```

🙀

Jaeyoung Chun

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 ...
```

Jaeyoung Chun

디자인 O

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

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

Ⓐ

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

Ⓑ

customerNumber	orderCount
129	3
131	1

Jaeyoung Chun

디자인 1

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 🙀

Jaeyoung Chun

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

자식
2023-03-24

부모129

자식
2023-04-15

자식
2023-04-01
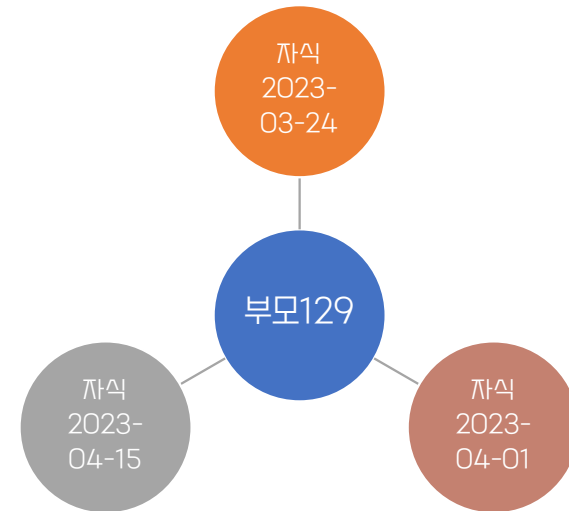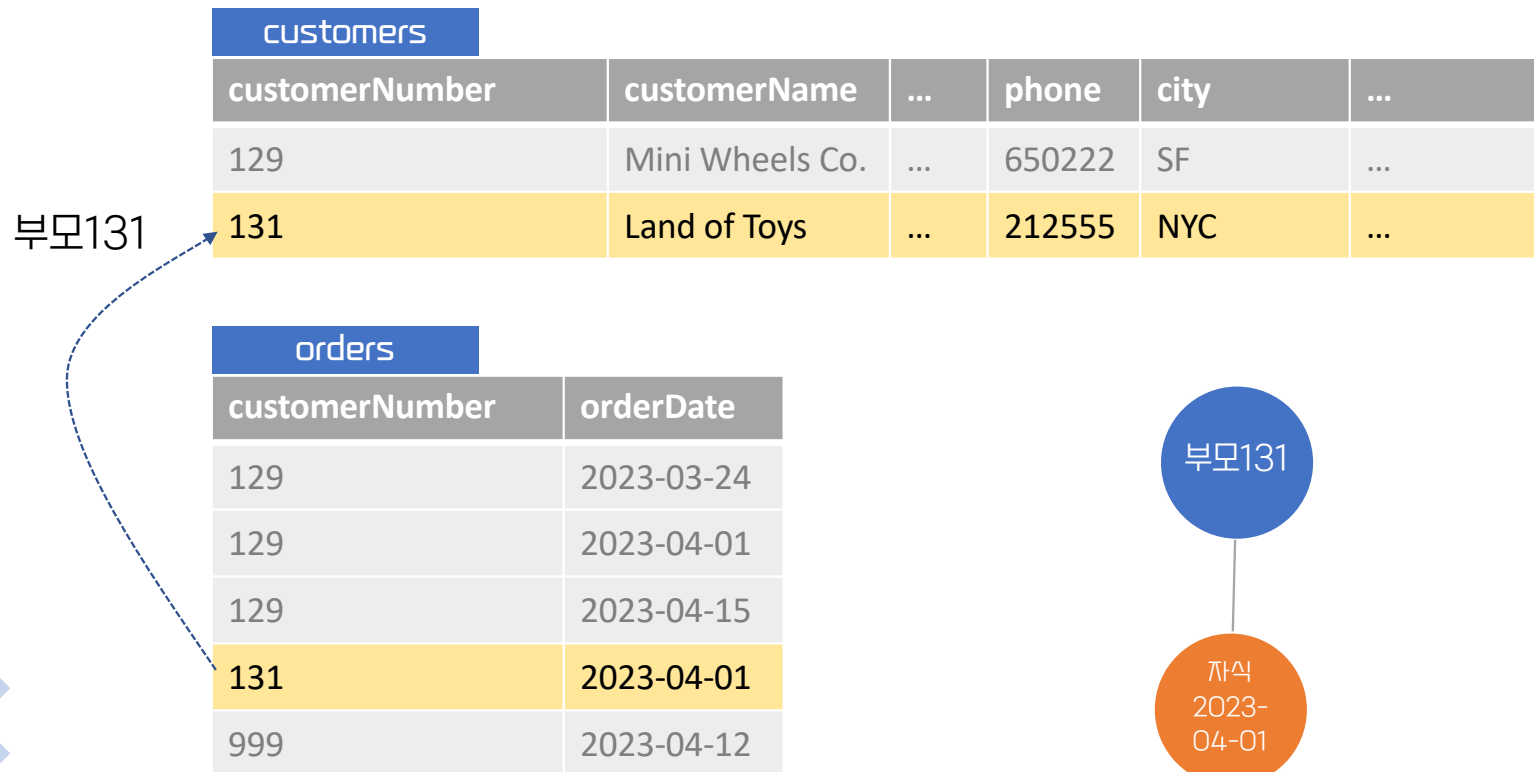
부모/자식

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?

자식
2023-
04-12

orphan record

고아 찾기

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.

Jaeyoung Chun

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

PK (Primary Key)

동일한 고백번호 131로 고객
테이블에 입력 불가능.

FK (Foreign Key)

고객번호 999로
주문입력 불가능.

Jaeyoung Chun

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?

Jaeyoung Chun

디자인 3

customers

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

orders

orderNumber PK	customerNumber FK	orderDate
100	129	2023-03-24
110	129	2023-04-01
120	129	2023-04-15
130	131	2023-04-01
140	129	2023-04-15

동일한 주문번호로 재입력 불가능

같은 고객, 같은 주문 날짜이지만,
주문번호 120

같은 고객, 같은 주문 날짜이지만,
주문번호 140

Jaeyoung Chun

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?

🙀

Jaeyoung Chun

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
100	129	2023-03-24
110	129	2023-04-01
120	129	2023-04-15
130	131	2023-04-01
140	129	2023-04-15

orderdetails

orderNumber FK	product	qty
100	A	1
...	...	...
120	C	2
140	C	1
140	F	1

Jaeyoung Chun