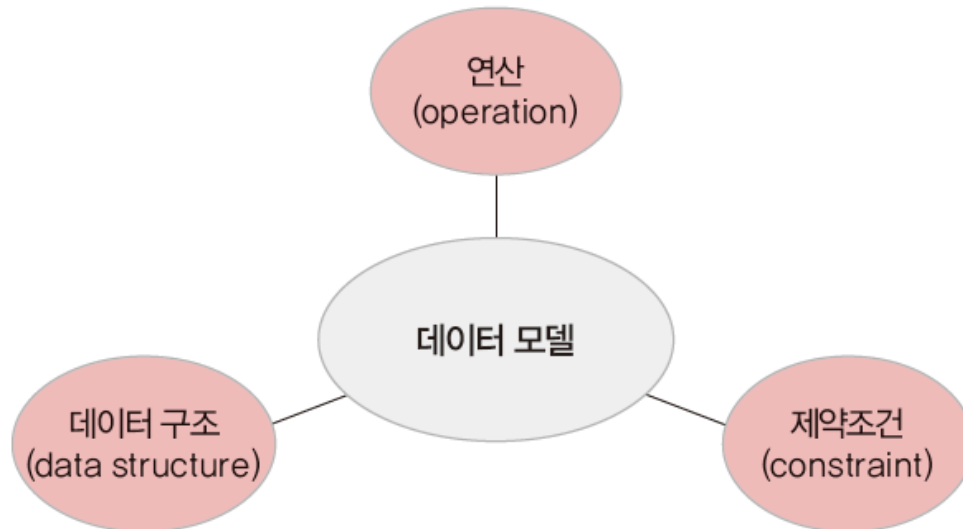


## 6장. 관계 데이터 연산

# 관계 데이터 연산의 개념

---

❖ 데이터 모델 = 데이터 구조 + 연산 + 제약조건



# 관계 데이터 연산의 개념

## ❖ 관계 데이터 연산(relational data operation)

- 릴레이션에 필요한 요구를 처리하는 것
- 관계 대수와 관계 해석이 있음
  - 기능과 표현력 측면에서 능력이 동등함



# 관계 대수

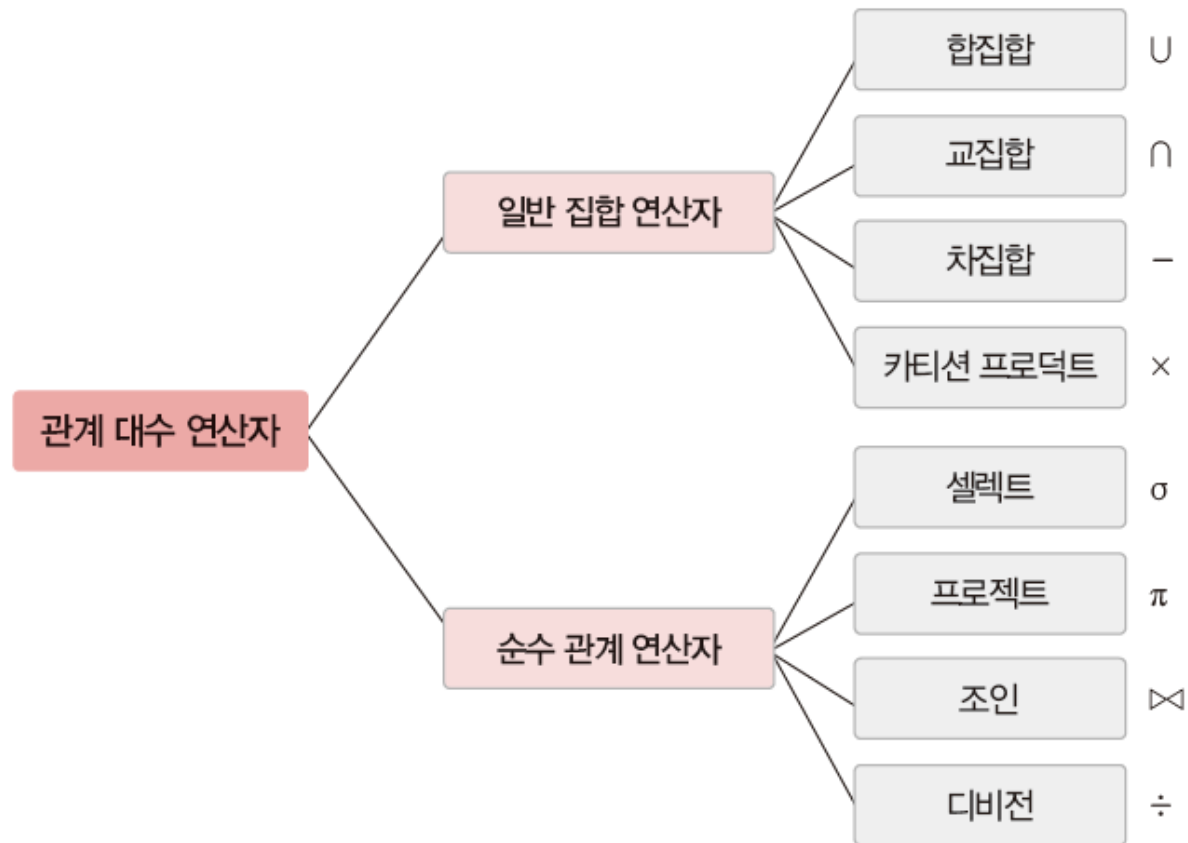
---

## ❖ 관계 대수(relational algebra)의 개념

- 원하는 결과를 얻기 위해 릴레이션의 처리 과정을 순서대로 기술하는 언어
  - 절차 언어(procedural language)
- 폐쇄 특성(closure property)
  - 피연산자도 릴레이션이고 연산의 결과도 릴레이션임

# 관계 대수

## ❖ 관계 대수의 연산자



# 관계 대수

## ❖ 일반 집합 연산자(set operation)

- 릴레이션이 튜플의 집합이라는 개념을 이용하는 연산자

연산자	기호	표현	의미
합집합	$\cup$	$R \cup S$	릴레이션 R과 S의 합집합을 반환
교집합	$\cap$	$R \cap S$	릴레이션 R과 S의 교집합을 반환
차집합	$-$	$R - S$	릴레이션 R과 S의 차집합을 반환
카티션 프로덕트	$\times$	$R \times S$	릴레이션 R의 각 튜플과 릴레이션 S의 각 튜플을 모두 연결하여 만들어진 새로운 튜플을 반환

# 관계 대수

---

## ❖ 일반 집합 연산자의 특성

- 합집합, 교집합, 차집합은 피연산자인 두 릴레이션이 합병 가능해야 함
  - 합병 가능(union-compatible) 조건
    - 두 릴레이션의 차수가 같아야 함
    - 두 릴레이션에서 서로 대응되는 속성의 도메인이 같아야 함

# 관계 대수

고객 릴레이션

고객번호	고객이름	나이
INT	CHAR(20)	INT
100	정소화	20
200	김선우	35
300	고명석	24

직원 릴레이션

직원번호	직원이름	직위
INT	CHAR(20)	CHAR(20)
10	김용욱	부장
20	채광주	과장
30	김수진	대리

고객 릴레이션

고객번호	고객이름	나이
INT	CHAR(20)	INT
100	정소화	20
200	김선우	35
300	고명석	24

직원 릴레이션

직원번호	직원이름	나이
INT	CHAR(20)	INT
10	김용욱	40
20	채광주	32
30	김수진	28



# 관계 대수

---

## ❖ 일반 집합 연산자 – 합집합(union)

- 합병 가능한 두 릴레이션 R과 S의 합집합 :  $R \cup S$ 
  - 릴레이션 R에 속하거나 릴레이션 S에 속하는 모든 튜플로 결과 릴레이션 구성
- 결과 릴레이션의 특성
  - 차수 : 릴레이션 R과 S의 차수와 같음
  - 카디널리티 : 릴레이션 R과 S의 카디널리티를 더한 것과 같거나 적어짐
- 교환법칙 성립 :  $R \cup S = S \cup R$
- 결합법칙 성립 :  $(R \cup S) \cup T = R \cup (S \cup T)$

# 관계 대수

---

R

번호	이름
100	정소화
200	김선우
300	고명석

S

번호	이름
100	정소화
101	채광주
102	김수진



합집합 연산

RUS

번호	이름
100	정소화
200	김선우
300	고명석
101	채광주
102	김수진

# 관계 대수

---

## ❖ 일반 집합 연산자 – 교집합(intersection)

- 합병 가능한 두 릴레이션 R과 S의 교집합 :  $R \cap S$ 
  - 릴레이션 R과 S에 공통으로 속하는 튜플로 결과 릴레이션 구성
- 결과 릴레이션의 특성
  - 차수 : 릴레이션 R과 S의 차수와 같음
  - 카디널리티 : 릴레이션 R과 S의 어떤 카디널리티보다 크지 않음
- 교환법칙 성립 :  $R \cap S = S \cap R$
- 결합법칙 성립 :  $(R \cap S) \cap T = R \cap (S \cap T)$

# 관계 대수

---

R

번호	이름
100	정소화
200	김선우
300	고명석

S

번호	이름
100	정소화
101	채광주
102	김수진

↓ 교집합 연산

$R \cap S$

번호	이름
100	정소화

# 관계 대수

---

## ❖ 일반 집합 연산자 – 차집합(difference)

- 합병 가능한 두 릴레이션 R과 S의 차집합 :  $R-S$ 
  - 릴레이션 R에는 존재하지만 릴레이션 S에는 존재하지 않는 튜플로 결과 릴레이션 구성
- 결과 릴레이션의 특성
  - 차수 : 릴레이션 R과 S의 차수와 같음
  - $R-S$ 의 카디널리티 : 릴레이션 R의 카디널리티와 같거나 적음
  - $S-R$ 의 카디널리티 : 릴레이션 S의 카디널리티와 같거나 적음
- 교환법칙과 결합법칙 모두 성립하지 않음

# 관계 대수

---

R

번호	이름
100	정소화
200	김선우
300	고명석

S

번호	이름
100	정소화
101	채광주
102	김수진

차집합 연산  
↓

R-S

번호	이름
200	김선우
300	고명석

S-R

번호	이름
101	채광주
102	김수진

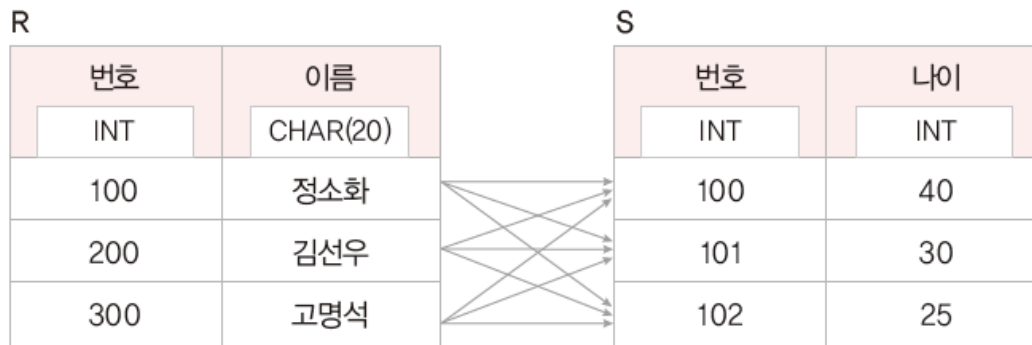
# 관계 대수

---

## ❖ 일반 집합 연산자 – 카티션 프로덕트(cartesian product)

- 두 릴레이션 R과 S의 카티션 프로덕트 :  $R \times S$ 
  - 릴레이션 R에 속한 각 튜플과 릴레이션 S에 속한 각 튜플을 모두 연결하여 만들어진 새로운 튜플로 결과 릴레이션을 구성
  - $R \times S = \{ r \cdot s \mid r \in R \wedge s \in S \}$ ,  $\cdot$  : 접속(concatenation)
- 결과 릴레이션의 특성
  - 차수 : 릴레이션 R과 S의 차수를 더한 것과 같음
  - 카디널리티 : 릴레이션 R과 S의 카디널리티를 곱한 것과 같음
- 교환법칙 성립 :  $R \times S = S \times R$
- 결합법칙 성립 :  $(R \times S) \times T = R \times (S \times T)$

# 관계 대수



카티션 프로덕트 연산

R × S

R.번호	R.이름	S.번호	S.나이
INT	CHAR(20)	INT	INT
100	정소화	100	40
100	정소화	101	30
100	정소화	102	25
200	김선우	100	40
200	김선우	101	30
200	김선우	102	25
300	고명석	100	40
300	고명석	101	30
300	고명석	102	25

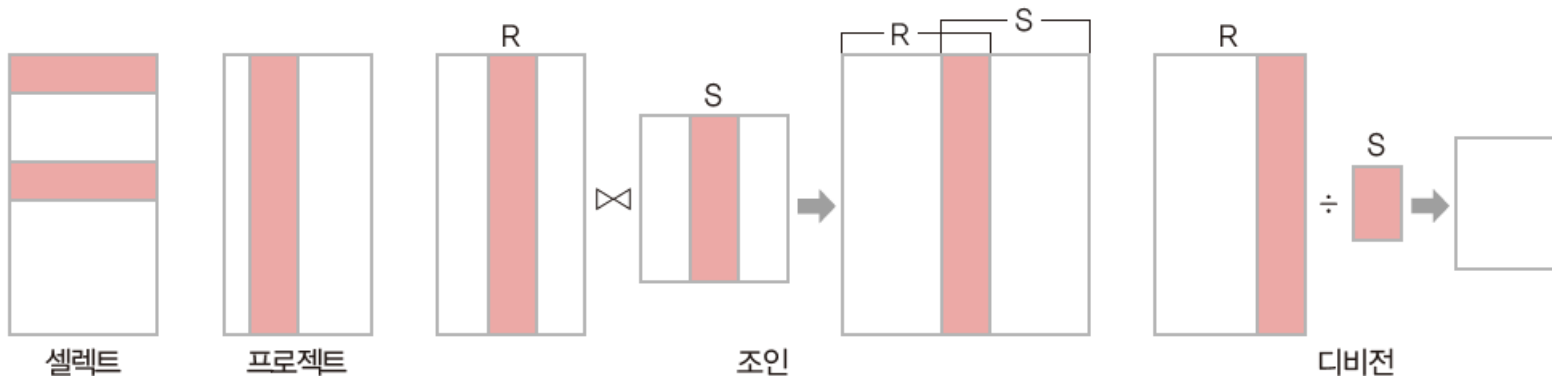


# 관계 대수

## ❖ 순수 관계 연산자(relational operation)

- 릴레이션의 구조와 특성을 이용하는 연산자

연산자	기호	표현	의미
선택	$\sigma$	$\sigma_{\text{조건}}(R)$	릴레이션 R에서 조건을 만족하는 튜플들을 반환
프로젝트	$\pi$	$\pi_{\text{속성리스트}}(R)$	릴레이션 R에서 주어진 속성들의 값으로만 구성된 튜플들을 반환
조인	$\bowtie$	$R \bowtie S$	공통 속성을 이용해 릴레이션 R과 S의 튜플들을 연결하여 만들어진 새로운 튜플들을 반환
디비전	$\div$	$R \div S$	릴레이션 S의 모든 튜플과 관련이 있는 릴레이션 R의 튜플들을 반환



# 관계 대수

---

## ❖ 순수 관계 연산자 – 선택(select)

### ■ 기능

- 릴레이션에서 조건을 만족하는 튜플만 선택하여 결과 릴레이션을 구성

### ■ 대상

- 하나의 릴레이션

### ■ 결과

- 연산 대상 릴레이션의 수평적 부분집합

### ■ 표현법 : $\sigma_{\text{조건식}}$ (릴레이션)

#### • 조건식

- 속성과 상수의 비교나 속성들 간의 비교로 표현
- 비교 연산자(>, ≥, <, ≤, =, ≠)와 논리 연산자(∧, ∨, ¬)를 이용해 작성

# 관계 대수

## ❖ 순수 관계 연산자 – 선택

고객아이디	고객이름	나이	등급	직업	적립금
apple	김현준	20	gold	학생	1000
banana	정소화	25	vip	간호사	2500
carrot	원유선	28	gold	교사	4500
orange	정지영	22	silver	학생	0

### 예제 6-1

고객 릴레이션에서 등급이 gold인 튜플을 검색하시오.

▶▶  $\sigma_{\text{등급='gold'}}(\text{고객})$

결과 릴레이션

고객아이디	고객이름	나이	등급	직업	적립금
apple	김현준	20	gold	학생	1000
carrot	원유선	28	gold	교사	4500

# 관계 대수

## ❖ 순수 관계 연산자 – 선택

고객아이디	고객이름	나이	등급	직업	적립금
apple	김현준	20	gold	학생	1000
banana	정소화	25	vip	간호사	2500
carrot	원유선	28	gold	교사	4500
orange	정지영	22	silver	학생	0

### 예제 6-2

고객 릴레이션에서 등급이 gold이고, 적립금이 2000 이상인 튜플을 검색하시오.

▶▶  $\sigma_{\text{등급}='gold' \wedge \text{적립금} \geq 2000}(\text{고객})$

결과 릴레이션

고객아이디	고객이름	나이	등급	직업	적립금
carrot	원유선	28	gold	교사	4500

# 관계 대수

---

## ❖ 순수 관계 연산자 – 선택

### ■ 교환법칙 성립

$$\sigma_{조건식1}(\sigma_{조건식2}(릴레이션)) = \sigma_{조건식2}(\sigma_{조건식1}(릴레이션)) = \sigma_{조건식1 \wedge 조건식2}(릴레이션)$$

$$\sigma_{적립금 \geq 2000}(\sigma_{등급='gold'}(고객)) = \sigma_{등급='gold'}(\sigma_{적립금 \geq 2000}(고객)) = \sigma_{등급='gold' \wedge 적립금 \geq 2000}(고객)$$

# 관계 대수

---

## ❖ 순수 관계 연산자 – 프로젝트(project)

- 기능
  - 릴레이션에서 선택한 속성의 값으로 결과 릴레이션을 구성
- 대상
  - 하나의 릴레이션
- 결과
  - 연산 대상 릴레이션의 수직적 부분집합
- 표현법 :  $\pi_{\text{속성리스트}}(\text{릴레이션})$

# 관계 대수

## ❖ 순수 관계 연산자 – 프로젝트

고객아이디	고객이름	나이	등급	직업	적립금
apple	김현준	20	gold	학생	1000
banana	정소화	25	vip	간호사	2500
carrot	원유선	28	gold	교사	4500
orange	정지영	22	silver	학생	0

### 예제 6-3

고객 릴레이션에서 고객이름, 등급, 적립금을 검색하시오.

▶▶  $\pi_{\text{고객이름, 등급, 적립금}}(\text{고객})$

결과 릴레이션

고객이름	등급	적립금
정소화	gold	1000
김선우	vip	2500
고명석	gold	4500
김용욱	silver	0

# 관계 대수

## ❖ 순수 관계 연산자 – 프로젝트

고객아이디	고객이름	나이	등급	직업	적립금
apple	김현준	20	gold	학생	1000
banana	정소화	25	vip	간호사	2500
carrot	원유선	28	gold	교사	4500
orange	정지영	22	silver	학생	0

### 예제 6-4

고객 릴레이션에서 등급을 검색하시오.

▶▶  $\pi_{\text{등급}}(\text{고객})$

결과 릴레이션

등급
gold
vip
silver

결과 릴레이션에서 동일한 튜플은 중복되지 않고 한 번만 나타남



# 관계 대수

---

## ❖ 순수 관계 연산자 – 세타 조인(theta join, $\theta$ -join)

### ■ 기능

- 주어진 조인 조건을 만족하는 두 릴레이션의 모든 튜플을 연결하여 생성된 새로운 튜플로 결과 릴레이션을 구성

- $R \bowtie_{A\theta B} S = \{ r \cdot s \mid r \in R \wedge s \in S \wedge (r.A \theta s.B) \}$

- A, B : 조인 속성

- $\theta$ 는 비교 연산자(>, ≥, <, ≤, =, ≠)를 의미

- 결과 릴레이션의 차수 : 두 릴레이션의 차수를 더한 것과 같음

## ❖ 동일 조인(equi-join)

- $\theta$  연산자가 “=”인 세타 조인을 의미

# 관계 대수

## ❖ 동일 조인

고객 릴레이션

고객아이디	고객이름	나이	등급
apple	김현준	20	gold
banana	정소화	25	vip
carrot	원유선	28	gold
orange	정지영	22	silver

주문 릴레이션

주문번호	주문고객	주문제품	수량
1001	apple	진짜우동	10
1002	carrot	맛있는파이	5
1003	banana	그대로만두	11



동일 조인 연산

고객 ⋈ 고객아이디=주문고객 주문

고객아이디	고객이름	나이	등급	주문번호	주문고객	주문제품	수량
apple	김현준	20	gold	1001	apple	진짜우동	10
banana	정소화	25	vip	1003	banana	그대로만두	11
carrot	원유선	28	gold	1002	carrot	맛있는파이	5

# 관계 대수

---

## ❖ 순수 관계 연산자 – 조인(join) 혹은 자연조인

### ■ 기능

- 조인 속성을 이용해 두 릴레이션을 조합하여 결과 릴레이션을 구성
  - 조인 속성의 값이 같은 튜플만 연결하여 생성된 튜플을 결과 릴레이션에 포함
  - 조인 속성 : 두 릴레이션이 공통으로 가지고 있는 속성
  - 동일조인의 결과 릴레이션에서 중복 속성을 제거한 결과와 동일함
- $R \bowtie_N S = \pi_{X \cup Y}(\sigma_{Z=Z}(R \times S)) = \pi_{X \cup Y}(R \bowtie_{Z=Z} S)$ 
  - $R(X)$ ,  $S(Y)$ 의 조인 속성 :  $Z(=X \cap Y)$

### ■ 대상

- 두 개의 릴레이션

# 관계 대수

## ❖ 순수 관계 연산자 – 조인

고객 릴레이션

고객아이디	고객이름	나이	등급
apple	김현준	20	gold
banana	정소화	25	vip
carrot	원유선	28	gold
orange	정지영	22	silver

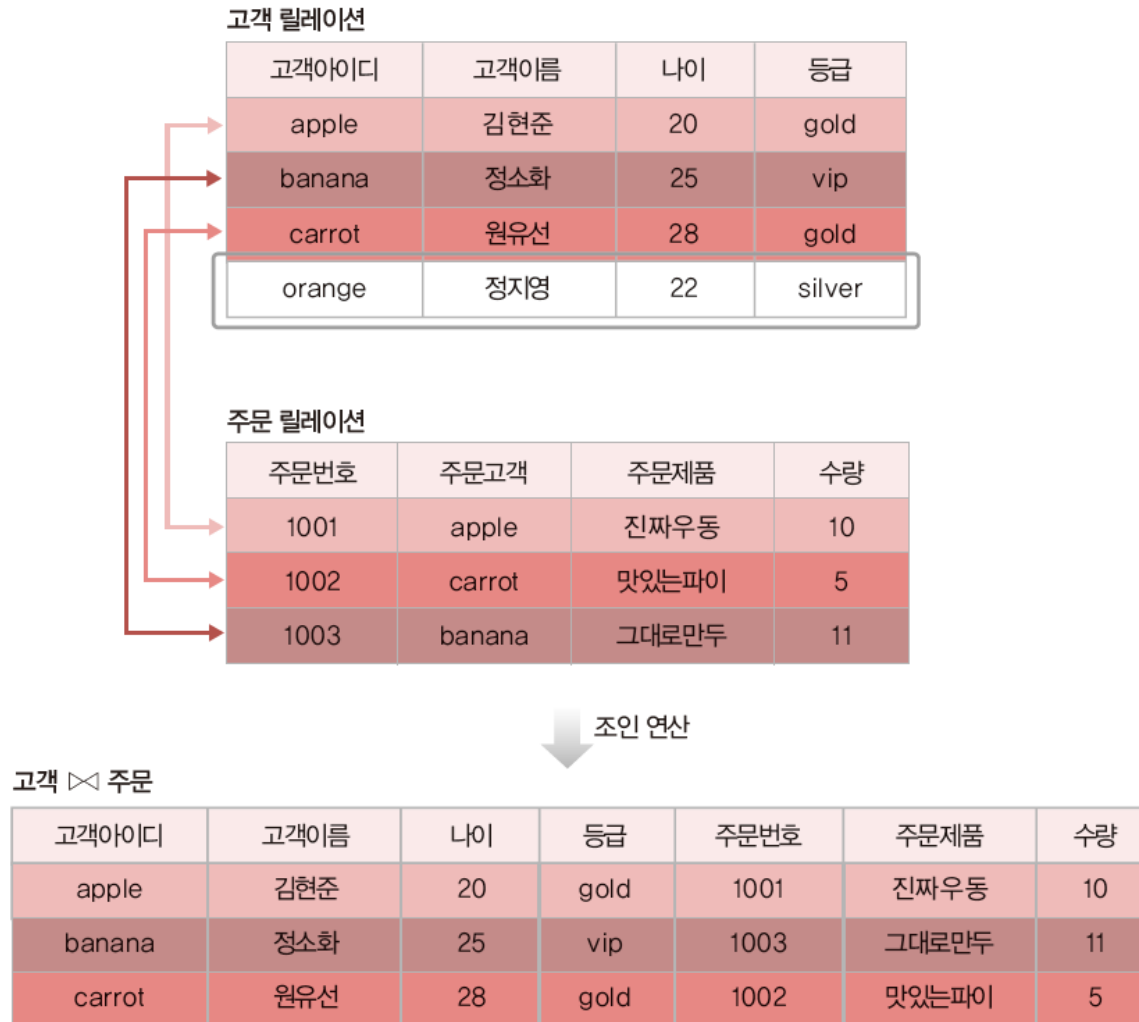
주문 릴레이션

주문번호	주문고객	주문제품	수량
1001	apple	진짜우동	10
1002	carrot	맛있는파이	5
1003	banana	그대로만두	11

주문 릴레이션의 외래키

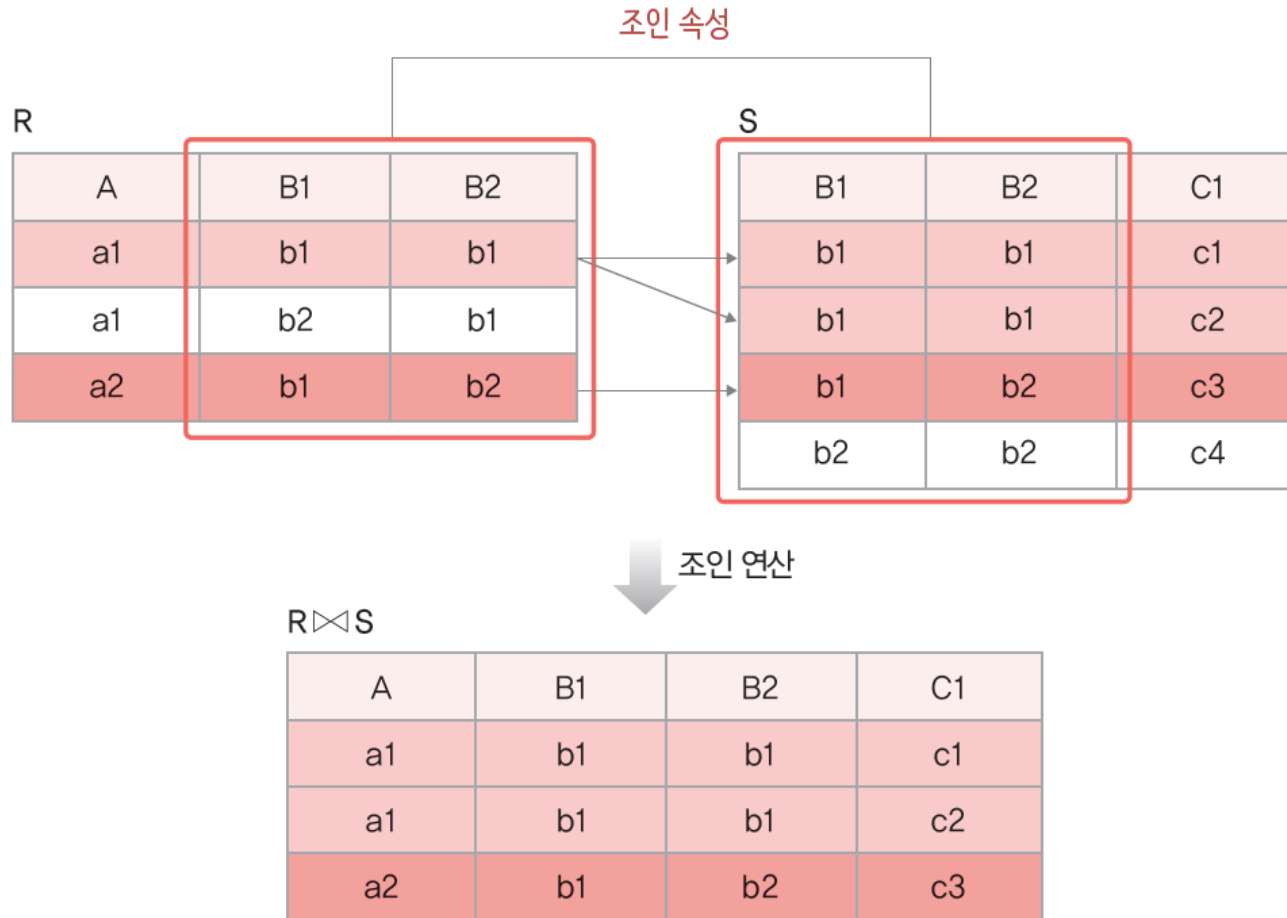
# 관계 대수

## ❖ 순수 관계 연산자 – 조인



# 관계 대수

## ❖ 순수 관계 연산자 – 조인



# 관계 대수

---

## ❖ 순수 관계 연산자 – 디비전(division)

- 릴레이션2의 모든 튜플과 관련이 있는 릴레이션1의 튜플로 결과 릴레이션을 구성
  - 단, 릴레이션1이 릴레이션2의 모든 속성을 포함하고 있어야 연산이 가능함
    - 도메인이 같아야 한다는 의미임
- 표현법 :  $R \div S$


$$12 \div 2 = (6 \times 2) \div 2 = 6$$

# 관계 대수

## ❖ 순수 관계 연산자 – 디비전

고객 릴레이션

고객아이디	고객이름	나이	등급	직업	적립금
apple	김현준	20	gold	학생	1000
NULL	정소화	25	vip	간호사	2500
carrot	원유선	28	gold	교사	4500
NULL	정지영	22	silver	학생	0

우수등급 릴레이션

등급
gold

↓ 디비전 연산

고객 ÷ 우수등급

고객아이디	고객이름	나이	직업	적립금
apple	김현준	20	학생	1000
carrot	원유선	28	교사	4500



# 관계 대수

## ❖ 순수 관계 연산자 – 디비전

주문내역 릴레이션

주문고객	제품이름	제조업체
apple	진짜우동	한빛식품
carrot	맛있는파이	마포과자
banana	그대로만두	한빛식품
apple	그대로만두	한빛식품
carrot	그대로만두	한빛식품

제품1 릴레이션

제품이름
진짜우동
그대로만두

제품2 릴레이션

제품이름	제조업체
그대로만두	한빛식품

↓ 디비전 연산

주문내역 ÷ 제품1

주문고객	제조업체
apple	한빛식품

주문내역 ÷ 제품2

주문고객
banana
apple
carrot

# 관계 대수

## ❖ 관계 대수를 이용한 질의 표현 예

고객 릴레이션

<u>고객아이디</u>	고객이름	나이	등급
apple	김현준	20	gold
banana	정소화	25	vip
carrot	원유선	28	gold
orange	정지영	22	silver

주문 릴레이션

<u>주문번호</u>	주문고객	주문제품	수량
1001	apple	진짜우동	10
1002	carrot	맛있는파이	5
1003	banana	그대로만두	11

주문 릴레이션의 외래키

# 관계 대수

## ❖ 관계 대수를 이용한 질의 표현 예

### 예제 6-5

등급이 gold인 고객의 이름과 나이를 검색하시오.

▶▶  $\pi_{\text{고객이름}, \text{나이}}(\sigma_{\text{등급}='gold'}(\text{고객}))$

결과 릴레이션

고객이름	나이
김현준	20
원유선	28

# 관계 대수

## ❖ 관계 대수를 이용한 질의 표현 예

### 예제 6-6

고객이름이 원유선인 고객의 등급과, 원유선 고객이 주문한 주문제품, 수량을 검색하시오.

▶▶  $\pi_{\text{등급, 주문제품, 수량}}(\sigma_{\text{고객이름='원유선'}}(\text{고객} \bowtie \text{주문}))$

결과 릴레이션

등급	주문제품	수량
gold	맛있는파이	5

# 관계 대수

## ❖ 관계 대수를 이용한 질의 표현 예

### 예제 6-7

주문수량이 10개 미만인 주문 내역을 제외하고 검색하시오.

▶▶ 주문 - ( $\sigma_{\text{주문수량} < 10}(\text{주문})$ )

결과 릴레이션

주문번호	주문고객	주문제품	수량
1001	apple	진짜우동	10
1003	banana	그대로만두	11

# 관계 대수

## ❖ 확장된 관계 대수 연산자 – 세미 조인(semi-join)

- 조인 속성으로 프로젝트 연산을 수행한 릴레이션을 이용하는 조인
- 릴레이션2를 조인 속성으로 프로젝트 연산한 후, 릴레이션1에 자연 조인하여 결과 릴레이션을 구성
  - $R \ltimes S = R \bowtie_N (\pi_Z(S)) = \pi_X(R \bowtie_N S)$ 
    - $R(X), S(Y)$ 의 조인 속성 :  $Z(=X \cap Y)$
- 장점
  - 최종 조인에 불필요한 속성이 미리 제거됨
    - 전송되는 데이터 양 등 조인 연산 비용을 크게 줄일 수 있음
- 교환법칙 불성립
  - $R \ltimes S \neq S \ltimes R$
  - $R \bowtie_N S = \pi_{X \cup Y}(\sigma_{Z=Z}(R \times S)) = \pi_{X \cup Y}(R \bowtie_{Z=Z} S)$ 
    - $R(X), S(Y)$ 의 조인 속성 :  $Z(=X \cap Y)$
- $R \bowtie_N S = (R \ltimes S) \bowtie_N S = (S \ltimes R) \bowtie_N R$

# 관계 대수

고객 릴레이션

고객아이디	고객이름	나이
apple	김현준	20
banana	정소화	25
carrot	원유선	28
orange	정지영	22

주문 릴레이션

주문번호	주문고객	주문제품
1001	apple	진짜우동
1002	carrot	맛있는파이
1003	banana	그대로만두

조인 속성

# 관계 대수

고객 릴레이션 ( R )

고객아이디	고객이름	나이
apple	김현준	20
banana	정소화	25
carrot	원유선	28
orange	정지영	22

주문 릴레이션 ( S )

주문번호	주문고객	주문제품
1001	apple	진짜우동
1002	carrot	맛있는파이
1003	banana	그대로만두

↓ 자연 조인 연산 (  $R \bowtie_N S$  )

고객  $\bowtie$  주문

고객아이디	고객이름	나이	주문번호	주문제품
apple	김현준	20	1001	진짜우동
banana	정소화	25	1003	그대로만두
carrot	원유선	28	1002	맛있는파이



# 관계 대수

## ❖ 확장된 관계 대수 연산자 – 세미 조인

고객 릴레이션 (  $R$  )

고객아이디	고객이름	나이
apple	김현준	20
banana	정소화	25
carrot	원유선	28
orange	정지영	22

주문 릴레이션 (  $S$  )

주문번호	주문고객	주문제품
1001	apple	진짜우동
1002	carrot	맛있는파이
1003	banana	그대로만두

$\pi_{\text{주문고객(주문)}} (\pi_Z(S))$

주문고객
apple
carrot
banana

자연 조인 연산 (  $R \bowtie_N (\pi_Z(S))$  )

고객  $\bowtie$  주문 (  $R \bowtie S$  )

고객아이디	고객이름	나이
apple	김현준	20
banana	정소화	25
carrot	원유선	28

# 관계 대수

## ❖ 확장된 관계 대수 연산자 – 세미 조인

고객 ⋈ 주문 (  $R \bowtie S$  )

고객아이디	고객이름	나이
apple	김현준	20
banana	정소화	25
carrot	원유선	28

주문 릴레이션 (  $S$  )

주문번호	주문고객	주문제품
1001	apple	진짜우동
1002	carrot	맛있는파이
1003	banana	그대로만두

↓ 자연 조인 연산 (  $(R \bowtie S) \bowtie_N S$  )

고객 ⋈ 주문 (  $R \bowtie_N S$  )

고객아이디	고객이름	나이	주문번호	주문제품
apple	김현준	20	1001	진짜우동
banana	정소화	25	1003	그대로만두
carrot	원유선	28	1002	맛있는파이

# 관계 대수

---

## ❖ 확장된 관계 대수 연산자 – 외부 조인(outer-join)

- 자연 조인 연산에서 제외되는 튜플도 결과 릴레이션에 포함시키는 조인
  - 두 릴레이션에 있는 모든 튜플을 결과 릴레이션에 포함시킴
- 표현법 : 릴레이션1  $\bowtie^+$  릴레이션2

# 관계 대수

## ❖ 확장된 관계 대수 연산자 – 외부 조인

고객 릴레이션

고객아이디	고객이름	나이
apple	김현준	20
banana	정소화	25
carrot	원유선	28
orange	정지영	22

주문 릴레이션

주문번호	주문고객	주문제품
1001	apple	진짜우동
1002	carrot	맛있는파이
1003	banana	그대로만두

↓ 외부 조인 연산

고객 ⋈<sup>+</sup> 주문

고객아이디	고객이름	나이	주문번호	주문제품
apple	김현준	20	1001	진짜우동
banana	정소화	25	1003	그대로만두
carrot	원유선	28	1002	맛있는파이
orange	정지영	22	NULL	NULL