

랩세미나 4주차



SQLD

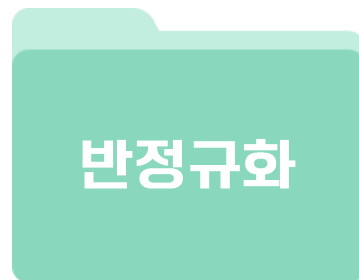
강한성 권우현 박준성

2025.02.13

목차



CONTENTS 01



CONTENTS 02



CONTENTS 03



CONTENTS 04



정규화

데이터의 정합성을 위해 엔티티를 작은 단위로 분리하는 과정

제1정규형

- 모든 속성은 반드시 하나의 값만 가져야 한다
- 유사한 속성이 반복되는 경우도 1차 정규화의 대상이 된다.

제2정규형

- 엔티티의 모든 일반 속성은 반드시 모든 주식별자에 종속되어야 한다.

제3정규형

- 주 식별자가 아닌 모든 속성 간에는 서로 종속될 수 없다.

제 1 정규형

모든 속성은 반드시 하나의 값만 가져야 한다

이름	생년월일	직업
이지은	19930516	배우, 가수, 작곡가
...



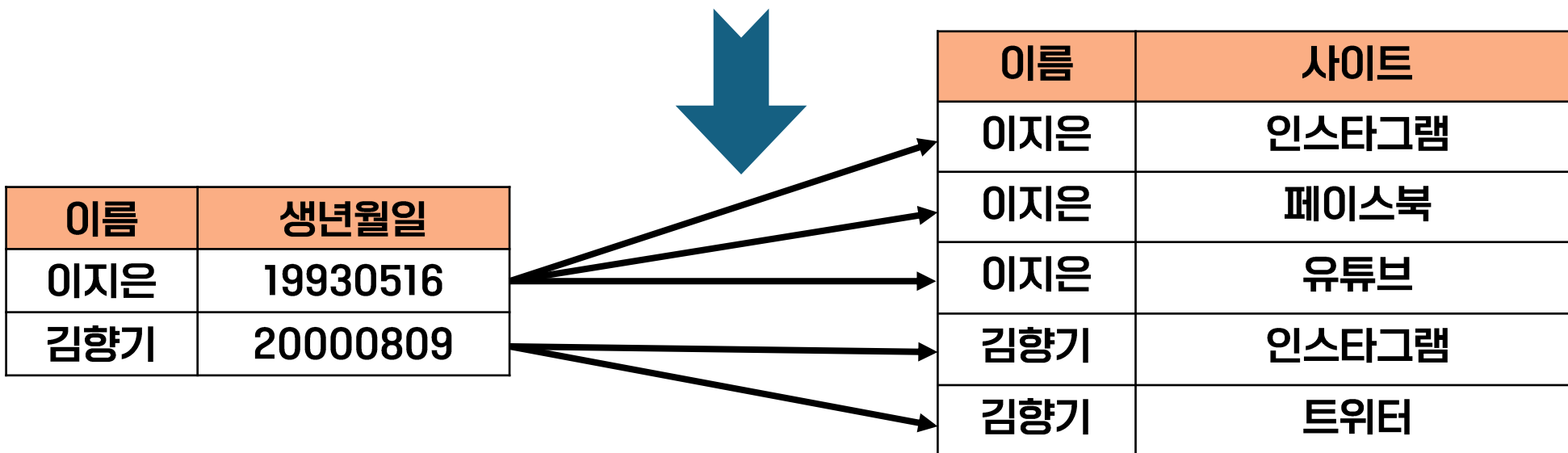
이름	생년월일
이지은	19930516

이름	직업
이지은	배우
이지은	가수
이지은	작곡가

제 1 정규형

유사한 속성이 반복되는 경우도 1차 정규화의 대상이 된다

이름	생년월일	사이트1	사이트2	사이트3
이지은	19930516	인스타그램	페이스북	유튜브
김향기	20000809	인스타그램	트위터	NULL
...



제 2 정규형

엔터티의 모든 일반 속성은 반드시 모든 주식별자에 종속되어야 한다.

주식별자

주문번호	음료코드	주문수량	음료명
202506010010	A1001	2	아메리카노
202506010011	B1002	1	카페라떼
202506010012	A1001	3	아메리카노

*일반속성인 음료명이 주식별자 중 음료코드 속성에만 종속된 경우

주문번호	음료코드	주문수량
202506010010	A1001	2
202506010011	B1002	1
202506010012	A1001	3

음료코드	음료명
A1001	아메리카노
B1002	카페라떼

제 3 정규형

주식별자가 아닌 모든 속성 간에는 서로 종속될 수 없다.

주식별자

일련번호	이름	생년월일	소속사 코드	소속사명
1	이지은	19930516	A1001	EDAM엔터테인먼트
2	김향기	20000809	B1004	지킴엔터테인먼트
3	유건우	19830925	B1004	지킴엔터테인먼트

*일반속성인 소속사명이 다른
일반속성인 소속사코드에 종속된 경우

일련번호	이름	생년월일	소속사 코드
1	이지은	19930516	A1001
2	김향기	20000809	B1004
3	유건우	19830925	B1004

소속사 코드	소속사명
A1001	EDAM엔터테인먼트
B1004	지킴엔터테인먼트

반정규화

- 데이터의 조화 성능을 향상시키기 위해 데이터의 중복을 허용하거나 데이터를 그룹핑하는 과정
- 조화성능은 향상될 수 있으나, 입력, 수정, 삭제 성능은 저하될 수 있으며, 데이터 합성 이슈가 발생할 수 있음
- 정규화가 끝난 후 거칠 수 있음

(1) 관계 반정규화(중복관계 추가)

- 업무 프로세스상 JOIN이 필요한 경우가 많아
중복 관계를 추가하는 것이 성능 측면에서 유리할 경우 고려한다.

테이블 반정규화

테이블 병합	1:1 관계 테이블 병합
	1:M 관계 테이블 병합
	슈퍼 서브 타입 테이블 병합
테이블 분할	테이블 수직 분할(속성 분할)
	테이블 수평 분할(인스턴스 분할, 파티셔닝)
테이블 추가	중복 테이블 추가
	통계 테이블 추가
	이력 테이블 추가
	부분 테이블 추가

컬럼 반정규화

1. 중복 컬럼 추가

- 업무 프로세스상 JOIN이 필요한 경우가 많아
컬럼을 추가하는 것이 성능 측면에서 유리할 경우 고려한다.

2. 파생 컬럼 추가

- 프로세스 수행 시 부하가 염려되는 계산 값을 컬럼으로 미리 추가하여 보관하는 방식
- 상품의 재고나 프로모션 적용 할인가 등이 이에 해당할 수 있다.

3. 이력 테이블 컬럼 추가

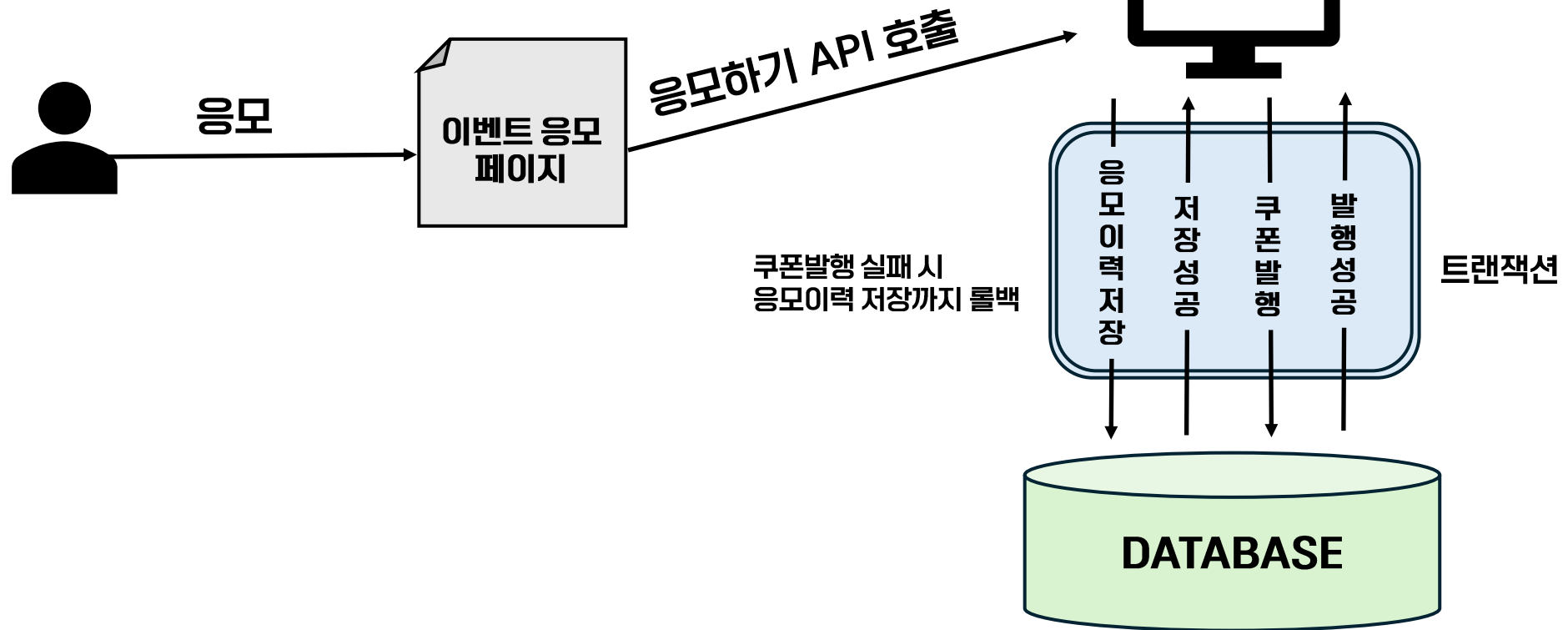
- 대량의 이력 테이블을 조회할 때 속도가 느려질 것을 대비하여
조회 기준이 될 것으로 판단되는 컬럼을 미리 추가해 놓는 방식이다.
- 최신 데이터 여부 등이 이에 해당할 수 있다.

• 트랜잭션(Transaction)

데이터를 조작하기 위한 하나의
논리적인 작업 단위

예시

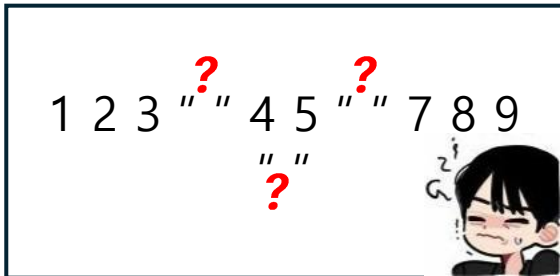
- 이벤트 응모 이력을 저장한다
- 쿠폰을 발행한다



NULL

NULL

- 존재하지 않음, 값이 없음



- NULL은 정의되지 않은 값으로 해석
- NULL과 다른 값은 비교할 수 없음
ex) NULL > 3 = Error
- NULL과 'NULL'은 같은 데이터가 아니다
*NULL은 값x, 'NULL'은 텍스트

(단위 : 만 원)

생년월일	이름	수입	지출	직업
20020101	박준성	0	2000	유튜버
19371231	강한성	50000	30	가공육제조
19080230	권우현	NULL	10	일식집 운영

위 데이터에서 0과 NULL값이 존재
둘의 의미는 엄연히 다름
박준성의 수입은 0원으로 확실하지만,
권우현의 수입은 0원 ~ 수천만원 등 스펙트럼이 넓음

-> 데이터가 입력되지 않았기에, NULL값으로 표기된 것

NULL

NULL의 연산

가로 연산

(단위 : 만 원)

생년월일	이름	수입	지출	직업
20020101	박준성	0	2000	유튜버
19371231	강한성	50000	30	가공육제조
19080230	권우현	NULL	10	일식집 운영

Result : NULL

가로 연산 시,
NULL이 포함되어 있으면
결과값도 NULL이다

세로 연산

(단위 : 만 원)

생년월일	이름	수입	지출	직업
20020101	박준성	0	2000	유튜버
19371231	강한성	50000	30	가공육제조
19080230	권우현	NULL	10	일식집 운영

Result : 50000

세로 연산 시,
NULL이 포함되어 있으면
NULL값은 제외한다

감사합니다

Thank you

