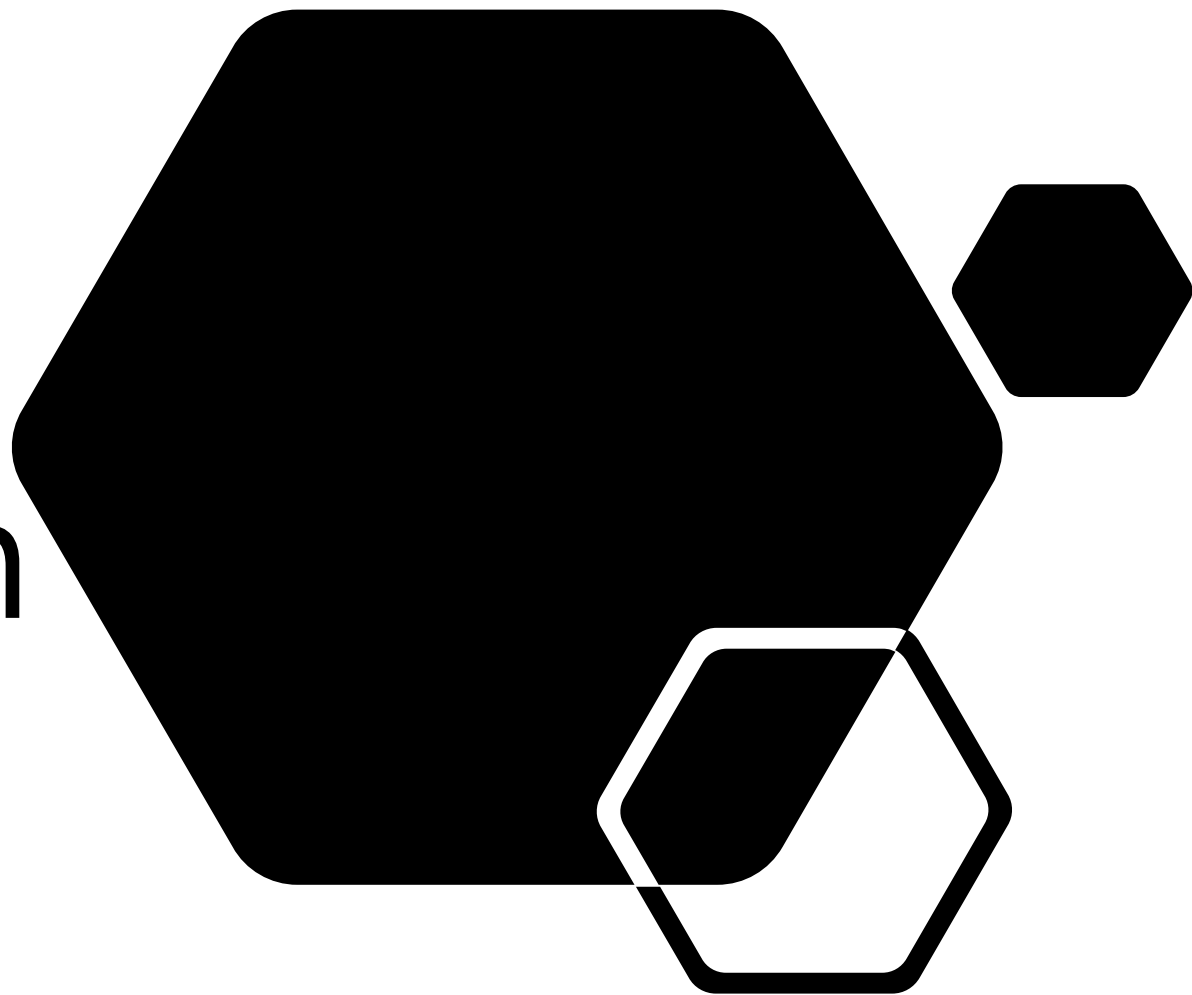


# Normalization

정규화



- 이상(anomaly) 현상

- 불필요한 데이터 중복으로 인해 릴레이션에 대한 데이터 삽입·수정·삭제 연산을 수행할 때 발생할 수 있는 부작용

삽입 이상	새 데이터를 삽입하기 위해 불필요한 데이터도 함께 삽입해야 하는 문제
갱신 이상	중복 튜플 중 일부만 변경하여 데이터가 불일치하게 되는 모순의 문제
삭제 이상	튜플을 삭제하면 꼭 필요한 데이터까지 함께 삭제되는 데이터 손실의 문제

- 이상(anomaly) 문제

- 이상 현상을 설명하기 위한 릴레이션의 예

<u>고객아이디</u>	<u>이벤트번호</u>	당첨여부	고객이름	등급
apple	E001	Y	정소화	gold
apple	E005	N	정소화	gold
apple	E010	Y	정소화	gold
banana	E002	N	김선우	vip
banana	E005	Y	김선우	vip
carrot	E003	Y	고명석	gold
carrot	E007	Y	고명석	gold
orange	E004	N	김용욱	silver

이벤트 참여 릴레이션

- 이상(anomaly) 문제

- 삽입 이상(insertion anomaly)

- 릴레이션에 새 데이터를 삽입하려면 불필요한 데이터도 함께 삽입해야 하는 문제

- [삽입 이상이 발생하는 예]

- 이벤트 참여 릴레이션은 삽입 이상이 발생함

- 아직 이벤트에 참여하지 않았고 아이디가 "melon", 이름이 "성원용", 등급이 "gold"인 신규 고객의 데이터는 이벤트참여 릴레이션에 삽입할 수 없음

- 삽입하려면 실제로 참여하지 않은 임시 이벤트번호를 함께 삽입해야 함

melon	NULL	NULL	성원용	gold	← 삽입 불가!
-------	------	------	-----	------	----------

- 이상(anomaly) 문제

- 갱신 이상(update anomaly)

- 릴레이션의 중복된 튜플들 중 일부만 수정하여 데이터가 불일치하게 되는 모순이 발생하는 문제

- [갱신 이상이 발생하는 예]

- 이벤트참여 릴레이션은 갱신 이상이 발생함

- 아이디가 "apple"인 고객의 등급이 "gold"에서 "vip"로 변경되었는데, 일부 튜플에 대해서만 등급이 수정된다면 "apple" 고객이 서로 다른 등급을 가지는 모순이 발생

고객아이디	이벤트번호	당첨여부	고객이름	등급
apple	E001	Y	정소화	vip
apple	E005	N	정소화	vip
apple	E010	Y	정소화	gold

← 데이터 불일치 발생!

- 이상(anomaly) 문제

- 삭제 이상(deletion anomaly)

- 릴레이션에서 튜플을 삭제하면 꼭 필요한 데이터까지 손실되는 연쇄 삭제 현상이 발생하는 문제

- [삭제 이상이 발생하는 예]

- 이벤트참여 릴레이션은 삭제 이상이 발생함

- 아이디가 "orange"인 고객이 이벤트 참여를 취소해 관련 튜플을 삭제하게 되면 이벤트 참여와 관련이 없는 고객아이디, 고객이름(김용욱), 등급(silver) 데이터까지 손실됨

orange	E004	N	김용욱	silver
--------	------	---	-----	--------

← 데이터 손실 발생!

- 정규화

- 이상 현상이 발생하지 않도록, 릴레이션을 관련 있는 속성들로만 구성하기 위해 릴레이션을 분해(decomposition)하는 과정
- 함수적 종속성을 판단해, 릴레이션을 연관성이 있는 속성들로만 구성되도록 분해하여 이상 현상이 발생하지 않는 올바른 릴레이션으로 만들어 나가는 과정

- 함수적 종속성(FD; Functional Dependency)

- 속성들 간의 관련성
- “X가 Y를 함수적으로 결정한다”
  - $X \rightarrow Y$ 로 표현(X는 결정자, Y는 종속자)

## ■ 함수 종속 관계 판단 예(1)

고객아이디	고객이름	등급
apple	정소화	gold
banana	김선우	vip
carrot	고명석	gold
orange	김용욱	silver

고객 릴레이션

- 속성 값은 계속 변할 수 있으므로 속성 자체의 특성과 의미를 기반으로 함수 종속성을 판단해야 함
- 일반적으로 기본키와 후보키는 릴레이션의 다른 모든 속성들을 함수적으로 결정함
- 기본키나 후보키가 아니어도 다른 속성 값을 유일하게 결정하는 속성은 함수 종속 관계에서 결정자가 될 수 있음

고객아이디 → 고객이름  
고객아이디 → 등급

또는

고객아이디 → (고객이름, 등급)

고객 릴레이션에 존재하는 함수적 종속 관계



## ■ 함수 종속 관계 판단 예(2)

고객아이디  $\rightarrow$  고객이름

{고객아이디, 이벤트번호}  $\rightarrow$  당첨여부

{고객아이디, 이벤트번호}  $\rightarrow$  고객이름

이벤트 릴레이션에 존재하는 함수적 종속 관계

고객이름은 {고객아이디, 이벤트번호}에 부분 함수 종속됨

- 완전 함수 종속(FFD; Full Functional Dependency)

- 릴레이션에서 속성 집합 Y가 속성 집합 X에 함수적으로 종속되어 있지만, 속성 집합 X의 전체가 아닌 일부분에는 종속되지 않음을 의미
- 일반적으로 함수 종속은 완전 함수 종속을 의미함
- 예) 당첨여부는 {고객아이디, 이벤트번호}에 완전 함수 종속됨

- 부분 함수 종속(PFD; Partial Functional Dependency)

- 릴레이션에서 속성 집합 Y가 속성 집합 X의 전체가 아닌 일부분에도 함수적으로 종속됨을 의미
- 예) 고객이름은 {고객아이디, 이벤트번호}에 부분 함수 종속됨



- 다음 릴레이션을 정규화 하세요

<u>고객아이디</u>	<u>이벤트번호</u>	당첨여부	등급	할인율
apple	E001, E005, E010	Y, N, Y	gold	10%
banana	E002, E005	N, Y	vip	20%
carrot	E003, E007	Y, Y	gold	10%
orange	E004	N	silver	5%

- 다음 릴레이션을 정규화 하세요

수강\_2

학번	과목번호	학점	학년	과목개설학과	학과장
s001	c001	A	1	컴퓨터	박유찬
s002	c001	B	3	컴퓨터	박유찬
s002	c002	C	3	경영	김철수
s003	c004	A	4	컴퓨터	박유찬
s004	c003	B	2	정보통신	강민애
s004	c003	C	2	정보통신	강민애
s002	c005	F	3	컴퓨터	박유찬