

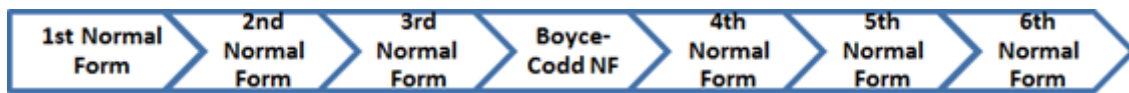
# 정규화(Normalization)

## 정규화란?

데이터의 무결성을 유지하기 위해

테이블 간 데이터의 **중복을 최소화**하고

**이상 현상을 제거**하는 데이터베이스 설계 기술



- 실제 애플리케이션에서는 3차 정규화에서 최상의 결과를 얻는다.
- 참고. 비정규화란?
  - 복잡한 쿼리 속도를 높이고 성능을 향상시키기 위해 테이블에 중복 데이터를 추가하는 프로세스

## 목적

- 불필요한 데이터를 최소화시킬 수 있다.
- 데이터가 논리적으로 저장되도록 한다.
- 데이터 삽입, 삭제, 갱신을 할 때 발생할 수 있는 이상현상들을 방지한다.
  - 삽입 이상 : 삽입 하려는 데이터가 아닌데 삽입이 된다거나, 삽입하려는데 데이터가 부족해 삽입이 되지 않는 문제
  - 삭제 이상 : 하나의 데이터를 삭제하고 싶는데 그 데이터가 포함된 튜플 전체가 삭제되어 원치않는 정보 손실이 발생하는 문제
  - 갱신 이상 : 일부 튜플만 갱신되어 정보가 정확하지 않거나 일관성이 없어지는 문제

## 장점

- 데이터베이스 변경시 이상 현상을 제거
- 데이터베이스 구조 확장시 재디자인을 최소화

## 단점

- 릴레이션 분해로 인해 릴레이션간의 연산(join)이 많아짐에 따라 응답 시간이 느려질 수 있다.

## 예시 테이블

- 대여한 영화 데이터베이스 관리
- 현재 정규화가 되지 않은 상태

FULL NAMES	PHYSICAL ADDRESS	MOVIES RENTED	SALUTATION
Janet Jones	First Street Plot No 4	Pirates of the Caribbean, Clash of the Titans	Ms.
Robert Phil	3 <sup>rd</sup> Street 34	Forgetting Sarah Marshal, Daddy's Little Girls	Mr.
Robert Phil	5 <sup>th</sup> Avenue	Clash of the Titans	Mr.

## 제 1정규화(1NF)

- 원자 값 구성
  - 각 테이블의 셀에는 하나의 값만 포함되어야한다.
  - 각 레코드(튜플)는 unique 해야 한다.

FULL NAMES	PHYSICAL ADDRESS	MOVIES RENTED	SALUTATION
Janet Jones	First Street Plot No 4	Pirates of the Caribbean	Ms.
Janet Jones	First Street Plot No 4	Clash of the Titans	Ms.
Robert Phil	3 <sup>rd</sup> Street 34	Forgetting Sarah Marshal	Mr.
Robert Phil	3 <sup>rd</sup> Street 34	Daddy's Little Girls	Mr.
Robert Phil	5 <sup>th</sup> Avenue	Clash of the Titans	Mr.

- 참고로 위 테이블에서는 복합 키를 기본키로 사용하고, 복합 키는 Full Names, Physical Address 이다.

## 제 2정규화(2NF)

- 먼저, 1NF에 속해야한다.
- 다음으로, **부분함수 종속 제거** 진행
  - 후보 키 관계의 하위 집합에 기능적으로 종속되지 않는 single column 기본 키를 만든다.
  - 기본키가 복합키일 때, 복합키 중 하나의 키만 가지고 다른 컬럼을 결정지을 수 있으면 안 된다. (부분 함수 종속)

MEMBERSHIP ID	FULL NAMES	PHYSICAL ADDRESS	SALUTATION
1	Janet Jones	First Street Plot No 4	Ms.
2	Robert Phil	3 <sup>rd</sup> Street 34	Mr.
3	Robert Phil	5 <sup>th</sup> Avenue	Mr.

MEMBERSHIP ID	MOVIES RENTED
1	Pirates of the Caribbean
1	Clash of the Titans
2	Forgetting Sarah Marshal
2	Daddy's Little Girls
3	Clash of the Titans

- 1NF 테이블을 회원 정보 테이블과 대여한 영화의 정보를 담은 테이블 두 개로 나누고, Membership ID라는 새로운 기본 키 컬럼을 도입하여 고유하게 식별할 수 있다.
- 대여한 영화 정보를 담은 테이블의 Membership ID는 외래키이다.
  - 외래키 : 다른 테이블의 기본 키 참조하여 테이블 연결

## 제 3정규화(3NF)

- 먼저, 2NF에 속해야한다.
- 다음으로, **이행함수 종속 제거** 진행
  - 이행 함수 종속이란?

- X, Y, Z 속성이 있을 때,  $X \rightarrow Y$ ,  $Y \rightarrow Z$ 란 종속 관계가 있다면  $X \rightarrow Z$ 도 성립되는 것.
- non-key 컬럼을 변경했을 때 다른 non-key 컬럼도 변경되는 것을 초래할 때 이행 함수 종속이 성립된다.
  - 2NF 한 테이블에서 Full Names 컬럼을 변경하면 Salutation이 변경될 수 있다.

MEMBERSHIP ID	FULL NAMES	PHYSICAL ADDRESS	SALUTATION ID
1	Janet Jones	First Street Plot No 4	2
2	Robert Phil	3 <sup>rd</sup> Street 34	1
3	Robert Phil	5 <sup>th</sup> Avenue	1

MEMBERSHIP ID	MOVIES RENTED
1	Pirates of the Caribbean
1	Clash of the Titans
2	Forgetting Sarah Marshal
2	Daddy's Little Girls
3	Clash of the Titans

SALUTATION ID	SALUTATION
1	Mr.
2	Ms.
3	Mrs.
4	Dr.

- 다시 테이블을 나누어 Salutation을 저장하는 새로운 테이블을 만듦
- 이행 함수 종속이 제거되었으므로 3NF를 만족한다.

## BCNF 정규화(3.5NF)

- 결정자 함수 종속
  - 데이터베이스가 3NF인 경우에도 하나 이상의 후보키가 있다면 이상 현상이 발생할 수 있다.
  - 3NF를 만족하면서 모든 결정자가 후보키 집합에 속할 때 BCNF를 만족한다.

## **제 4정규화(4NF)**

- 다치 종속성 제거

## **제 5정규화(5NF)**

- 조인 종속성 제거

## **Reference**

- <https://www.guru99.com/database-normalization.html>