

데이터 정규화란 중복데이터를 제거해서 데이터 무결성을 지키고 데이터를 효율적으로 저장하기 위한 일련의 규칙이다. 제1-5정규형, BCNF의 6가지 형태가 있다.

제1정규형은 컬럼이 원자값을 갖도록 하는 것이다.

제2정규형은 제1정규형을 만족하고 기본키가 아닌 모든 속성이 기본키를 구성하는 전체 속성에 대해 종속되는 완전 함수 종속을 만족해야 한다.

제3정규형은 제2정규형을 만족하고 이행적 종속이 없어야 한다. 이행적 종속이란 어떤 속성이 직접적으로 종속되는 것이 아닌 한다리 거쳐서 종속되는 것을 의미한다. 즉,  $A \Rightarrow B$  ( $B$ 가  $A$ 에 종속)  $B \Rightarrow C$  ( $C$ 가  $B$ 에 종속)이며  $A \Rightarrow C$  ( $C$ 가  $A$ 에 종속)되는 경우를 말한다.

BCNF는 제3 정규형의 강화된 형태로써 모든 결정자가 후보키가 되어야 한다는 규칙이다. 다른 속성이 종속되어있는 속성이 결정자이다. 즉, 그것의 값이 변하면 다른 속성값도 변하게 하는 컬럼을 의미하고 앞선 예시에서  $A \Rightarrow B$ 인 경우  $A$ 가  $B$ 의 결정자이다.

제4정규형은 BCNF를 만족하고 다치 종속성을 제거한 형태이다. 다치 종속성은 하나의 속성이 두 개 이상의 속성에 독립적으로 종속될 때 발생한다.

제5정규형은 제4정규형을 만족하고 모든 조인 종속성이 후보 키에 의해 암시적으로 보장되는 형태를 의미한다.