What Is Normalization?

In this lesson, we will learn about the different normalization forms and why we use the normalization technique.

Normalization should be part of the database design process. However, it is difficult to separate the normalization process from the ER modeling process so the two techniques should be used concurrently.

We use the entity relation diagram (ERD) to provide the big picture, or macro view, of an organization's data requirements and operations. This is created through an iterative process that involves identifying relevant entities, their attributes, and their relationships.

The normalization procedure focuses on the characteristics of specific entities and represents the micro view of entities within the entity-relationship diagram.

What is normalization?

Normalization is the branch of relational theory that provides design insights. It is the process of determining how much redundancy exists in a table. The goals of normalization are to:

- Be able to characterize the level of redundancy in a relational schema.
- Provide mechanisms for transforming schemas in order to remove redundancy.

Normalization draws heavily on the theory of functional dependencies.

Normalization draws heavily on the theory of functional dependencies. Normalization theory defines six normal forms (NF). Each normal form involves a set of dependency properties that a schema must satisfy and each normal form gives guarantees about the presence and/or absence of update anomalies. This means that higher normal forms have less redundancy, and as a result, fewer update problems.

Normal forms

There are six normal forms, but we will only look at the first four, which are:

- First normal form (1NF)
- Second normal form (2NF)
- Third normal form (3NF)
- Boyce-Codd normal form (BCNF)

Practical uses of normal forms

Most practical design projects in commercial and governmental environments acquire existing designs of databases from previous designs, from designs in legacy models, etc. They are certainly interested in the fact that the designs are of good quality and are sustainable over long periods of time. Existing designs are evaluated by applying the tests for normal forms, and normalization is carried out in practice so that the resulting designs are of high quality and have minimal data redundancy.

In the next lesson, we will discuss the first normal form in detail.