CP2420 Exercise 7-5

Curtis Baggs

**What is Normalization?**

Normalization is a database design technique which organizes tables in a manner that reduces redundancy and dependency of data.

It divides larger tables to smaller tables and links them using relationships.

The inventor of the relational model Edgar Codd proposed the theory of normalization with the introduction of First Normal Form, and he continued to extend theory with Second and Third Normal Form. Later he joined with Raymond F. Boyce to develop the theory of Boyce-Codd Normal Form.

Theory of Data Normalization in SQL is still being developed further. For example, there are discussions even on 6th Normal Form. However, in most practical applications, normalization achieves its best in 3rd Normal Form.

**1NF (First Normal Form) Rules:**

* Each table cell should contain a single value.
* Each record needs to be unique.

**2NF (Second Normal Form) Rules:**

* Be in 1NF
* Single Column Primary Key

**3NF (Third Normal Form) Rules:**

* Be in 2NF
* Has no transitive functional dependencies

**Boyce-Codd Normal Form (BCNF):**

* Even when a database is in 3rd Normal Form, still there would be anomalies resulted if it has more than one Candidate Key.
* Sometimes is BCNF is also referred as 3.5 Normal Form.

**4NF (Fourth Normal Form) Rules:**

* If no database table instance contains two or more, independent and multivalued data describing the relevant entity, then it is in 4th Normal Form.

**5NF (Fifth Normal Form) Rules:**

* A table is in 5th Normal Form only if it is in 4NF and it cannot be decomposed into any number of smaller tables without loss of data.