**Normalisation**

Normalization is a technique for organizing data in a database. It is important that a database is normalized to minimize redundancy (duplicate data) and to ensure only related data is stored in each table. It also prevents any issues stemming from database modifications such as insertions, deletions, and updates. [[1]](#footnote-1)  Normalisation is a technique in database design that organises tables, to reduce redundancy and dependency of data. It is a process of dividing larger tables into smaller tables and creating relationship links. It is important when designing a database that the design is both efficient and allows for data to be stored logically.

Normalization is a systematic approach of decomposing tables to eliminate data redundancy (repetition) and undesirable characteristics like Insertion, Update and Deletion Anomalies. It is a multi-step process that puts data into tabular form, removing duplicated data from the relation tables.

Normalization is used for mainly two purposes,

* Eliminating redundant (useless) data.
* Ensuring data dependencies make sense i.e data is logically stored.[[2]](#footnote-2)

The stages of organization are called normal forms. Each

First Normal Form (1NF):

* Data is stored in tables with rows uniquely identified by a primary key
* Data within each table is stored in individual columns in its most reduced form
* There are no repeating groups

Second Normal Form (2NF):

* Everything from 1NF
* Only data that relates to a table’s primary key is stored in each table

Third Normal Form (3NF):

* Everything from 2NF
* There are no in-table dependencies between the columns in each table[[3]](#footnote-3)

Boyce and Codd Normal Form:

* is a higher version of the Third Normal form. This form deals with certain type of anomaly that is not handled by 3NF. A 3NF table which does not have multiple overlapping candidate keys is said to be in BCNF. For a table to be in BCNF, following conditions must be satisfied:
* R must be in 3rd Normal Form
* and, for each functional dependency ( X → Y ), X should be a super Key.

Fourth Normal Form (4NF)

A table is said to be in the Fourth Normal Form when:

1. it is in the Boyce-Codd Normal Form.
2. And, it doesn't have Multi-Valued Dependency.[[4]](#footnote-4)

This normalisation systematic approach will be used to examine the following database (consisting of one table, with the primary key = EmployeeID) that was designed to store the following information on a company’s employees and departments:

• Employee ID

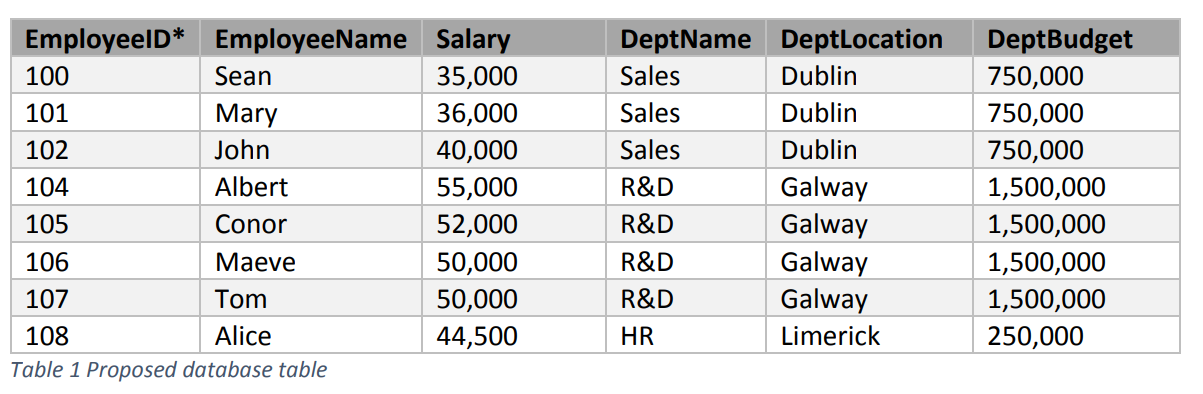
• Employee Name

• Employee Salary

• Department Name

• Department Location

• Department Budget



First Normal Form (1NF) Rules:

Each table cell should contain a single value – this is the case with the test database

Each record needs to be unique – this is true of the database. The primary key is the EmployeeID field, which is unique to each database record. It is important in this database construction to allow for a method that identifies that the primary key cannot by NULL and a primary key should be given a value when a new record is inserted. Ideally, the primary key should not be changed.

It would be a good if the test database had the addition of an employee surname or personnel number, which could act as a composite key.

2NF (Second Normal Form) Rules:

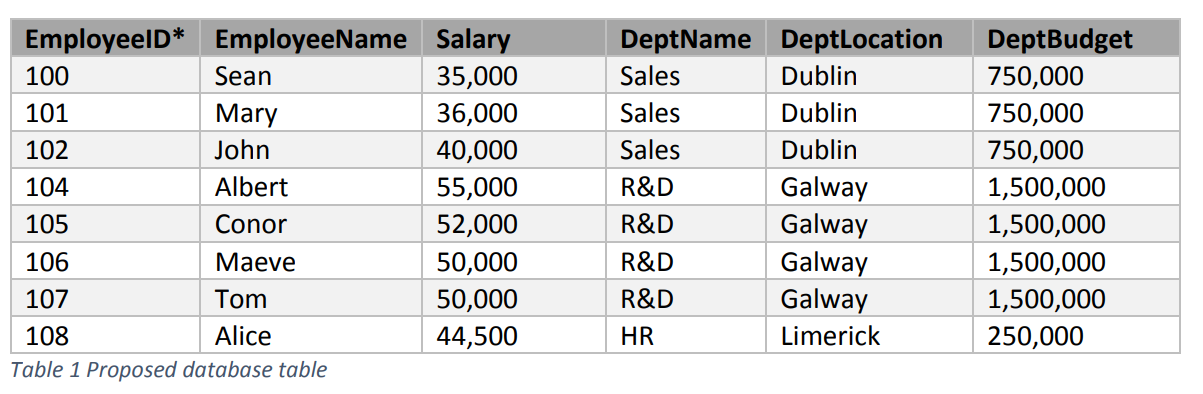
* To achieve second normal form, all data in each table must relate to the record that the primary key of the table identifies – this is achieved in the test database as the EmployeeID is the Primary Key.

3NF(Third Normal Form) Rules:

* To achieve third normal form, there cannot be transitive functional dependencies.

Transitively Dependent

Dependent Dependent



Unfortunately, there transitive dependencies in the test database, as shown in the above diagram.

1. <https://towardsdatascience.com/database-normalization-explained-53e60a494495> [↑](#footnote-ref-1)
2. <https://www.studytonight.com/dbms/database-normalization.php> [↑](#footnote-ref-2)
3. <https://towardsdatascience.com/database-normalization-explained-53e60a494495> [↑](#footnote-ref-3)
4. <https://www.studytonight.com/dbms/database-normalization.php> [↑](#footnote-ref-4)