**TODAY'S TOPIC: NORMALIZATION / DATA NORMALIZATION**

**Data:** Raw facts and figures.

**Information:** Organized for of data.

**Knowledge:** Practical side of information.

**Example:** Census

A **Table** is an arrangement of [data](https://en.wikipedia.org/wiki/Data) in **rows** and **columns**, and data in the Table may or may not be logically connected.

- may be confusing.

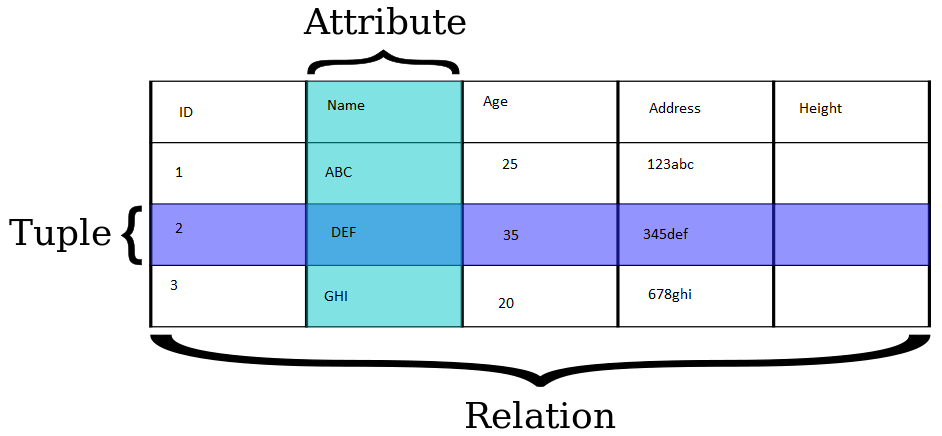
- may be repeating or may have redundant values.

- may be incomplete or having incomplete information.

- etc. etc.



A **Relation**, is a set of rows **(tuples)** and columns **(attributes)** which are logically connected to each other or having a predefined logical sequence. May also contains a unique number or identifier.



**- Unique ID: Uniquely Identified Number e.g. License #, Citizenship #, Contact # etc. etc.**

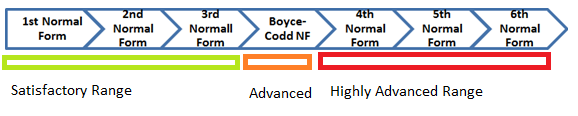
**- Strong ID vs. Weak ID**

**Database Normalization**, or simply **Normalization**, is the process of organizing the columns (attributes), and rows (tuples) in a **relation** (a specialized form of Table) in the database to reduce data redundancy and to improve data integrity/reliability.

**Forms of Normalization**

* First Normal Form **(1NF)**
* Second Normal Form **(2NF)**
* Third Normal Form **(3NF)**
* Boyce & Codd Normal Form **(BCNF)**
* Fourth Normal Form **(4NF)**
* Fifth Normal Form **(5NF)**
* Sixth Normal Form **(6NF)**

However, in most practical/conventional applications, normalization achieves its best in 3rd Normal Form. The evolution of Normalization theories is illustrated below;

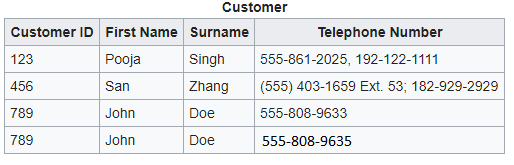


**First Normal Form** (**1NF**), A relation is in First (1st) Normal Form if and only if the [domain](https://en.wikipedia.org/wiki/Data_domain) of each [attribute](https://en.wikipedia.org/wiki/Column_(database)) contains **only**[**atomic**](https://en.wikipedia.org/wiki/First_normal_form#Atomicity) (indivisible) value, or the value of each attribute contains only a single value from that domain.

**A database is in 1NF if it satisfies the following conditions:**

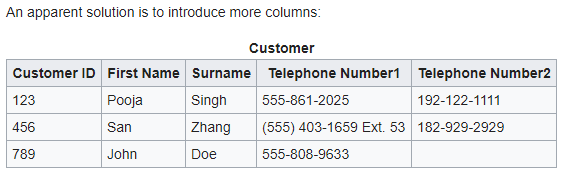
* Contains only atomic values
* There are no repeating groups in Parent Table

Example of a Table

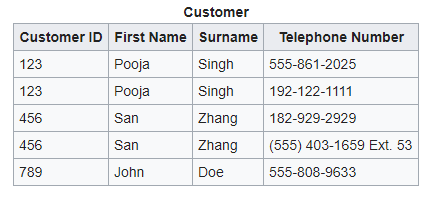


**What to do;**

* Create a separate Table for each set of related data.
* Identify each set of related data with a primary key.
* No repetition in Parent Table.



**Designs that is closer to 1NF**



**Designs that fully comply with 1NF**

