**What is Normalization?**Normalization is the process of efficiently organizing data in a database. There are two goals of the normalization process: eliminating redundant data (for example, storing the same data in more than one table) and ensuring data dependencies make sense (only storing related data in a table). Both of these are worthy goals as they reduce the amount of space a database consumes and ensure that data is logically stored.

**HEALTH HISTORY REPORT**

PET ID PET NAME PET TYPE PET AGE OWNER VISIT DATE PROCEDURE

246 ROVER DOG 12 SAM COOK JAN 13/2002 01 - RABIES VACCINATION

MAR 27/2002 10 - EXAMINE and TREAT WOUND

APR 02/2002 05 - HEART WORM TEST

298 SPOT DOG 2 TERRY KIM JAN 21/2002 08- TETANUS VACCINATION

MAR 10/2002 05 - HEART WORM TEST

341 MORRIS CAT 4 SAM COOK JAN 23/2001 01 - RABIES VACCINATION

JAN 13/2002 01 - RABIES VACCINATION

519 TWEEDY BIRD 2 TERRY KIM APR 30/2002 20 - ANNUAL CHECK UP

APR 30/2002 12 - EYE WASH

**First Normal Form (1NF)**

First normal form (1NF) sets the very basic rules for an organized database:

* Eliminate duplicative columns from the same table.
* Create separate tables for each group of related data and identify each row with a unique column or set of columns (the primary key).

**1NF:**Pet [ pet\_id, pet\_name, pet\_type, pet\_age, owner ]  
Pet\_Visit [ pet\_id, visitdate, procedure\_no, procedure\_name ]  
Note: a procedure may occur on multiple dates, therefore visitdate is included as part of the primary key. This is a composite key.

[**Selecting Primary Keys is an Extremely Important Decision**](http://databases.about.com/od/specificproducts/a/primarykey.htm)The selection of a primary key is one of the most critical decisions you’ll make in the design of a new database. The most important constraint is that you must ensure that the selected key is unique. If it’s possible that two records (past, present, or future) may share the same value for an attribute, it’s a poor choice for a primary key. You'll also need to avoid sensitive values, such as Social Security Numbers, as they raise privacy concerns.

## What is a composite Key?

* A composite key is a primary key composed of multiple columns used to identify a record uniquely.
* Use pet\_id and visitdate to uniquely identify a record. This is a composite key.**Second Normal Form (2NF)**

Second normal form (2NF) further addresses the concept of removing duplicative data:

* Meet all the requirements of the first normal form.
* Remove subsets of data that apply to multiple rows of a table and place them in separate tables.
* Create relationships between these new tables and their predecessors through the use of foreign keys.

Foreign Key references primary key of another Table. It helps connect your Tables.

* **Definition:** A foreign key is a field in a relational table that matches the primary key column of another table. The foreign key can be used to cross-reference tables.
* A foreign key can have a different name from its primary key
* It ensures rows in one table have corresponding rows in another
* Unlike Primary key they do not have to be unique. Most often they aren’t.
* Foreign keys can be null even though primary keys can not  
  <http://www.guru99.com/database-normalization.html#JH2W0k6L1dLamfdM.99>

**2NF:**

Pet [ pet\_id (PK), pet\_name, pet\_type, pet\_age, owner ]

Pet\_Visit [ pet\_id, visitdate (Composite Key), procedure\_no (FK) ]

Procedure [ procedure\_no (FK), procedure\_name ]

**Third Normal Form (3NF)**

Third normal form (3NF) goes one step further:

* Meet all the requirements of the second normal form.
* Remove columns that are not [dependent](http://databases.about.com/od/specificproducts/a/Database-Dependency.htm) upon the primary key.

**3NF:  
same as 2NF**

[**NULL Is Not Zero or the Empty String**](http://databases.about.com/cs/sql/a/aa042803a.htm)NULL is a very special value in the world of databases. When you see a NULL value, interpret it as "unknown". If a quantity is NULL, that doesn't necessarily mean that the quantity is zero. Similarly, if a text field holds a NULL value, that doesn't mean that there isn't an appropriate value, it's simply unknown. For example, consider a database containing information about children who attend a particular school. If the secretary entering the record does not know a student's age a NULL value is used to indicate the "unknown" placeholder. The student certainly has an age, it's just not present in the database.

**Resources:**<http://databases.about.com/od/specificproducts/a/normalization.htm>  
http://databases.about.com/od/administration/tp/beginners\_faq.htm