

Normalization of database structure

By: Robin Andersson, Simon Franzén

Index

3. Part one – Original tables
 - Candidate Keys found
 - Primary Keys found
 - Fully functional dependencies
4. - Highest form of normalization in tables
5. Part two – New construction of tables
 - First normalization
 - Second normalization
 - Third normalization
6. -The new tables
8. Discussion

Part one: The original tables

Candidate keys found:

<u>Table</u>	<u>CK</u>
Supplier	Number
Store	Number
Employee	Number
Debt	Number
Item	Number
Sale	Number + Item
Parts	Pnum
Supply	Snum + Pnum + Jnum

Primary keys found:

<u>Table</u>	<u>PK</u>
Supplier	Number
Store	Number
Employee	Number
Debt	Number
Item	Number
Sale	Number + Item
Parts	Pnum
Supply	Snum + Pnum + Jnum

Fully functional dependencies:

<u>Table</u>	<u>FFD</u>
Supplier	Number + name + city
Store	Number + city
Employee	Full row
Dept	Full row
Item	Full row
Sale	Number + Item -> Date + Quantity
Parts	Pnum
Supply	Snum + Pnum + Jnum -> Shipdate + Quantity

Highest form of normalization in tables

<u>Table</u>	<u>Form</u>
Supplier	2NF
Store	2NF
Employee	None*
Dept	BCNF
Item	BCNF
Sale	1NF
Parts	BCNF
Supply	BCNF

*Will be clarified in the discussion part of this paper

Part two: New constructions of tables

First normalization:

Table Employees:

The column “Name” is used by entering 2 names, first- and lastname, into the same field. By splitting up this column into two separate columns, “Firstname” and “Lastname”, searching for an employee will become easier.

Second normalization:

Table Sale:

There is a FD between Item and Dept, since a specific item will only ever be in one type of department.

There is also a FD between Number and Credit, where Credit is dependent on Number.

Third normalization:

Table Supplier:

Since there is a FD between the columns City and State, where state is dependent on City, the table Store won't qualify for the 3NF.

Table Store:

Since there is a FD between the columns City and State, where state is dependent on City, the table Store won't qualify for the 3NF.

New construction of tables

The tables “Supplier” and “Store” both had problems with redundancy caused of the FD “City”-“State”.

The changes below is made to correct this problem.

Supplier

<u>Number</u>	Name	City
199	Koret	Los Angeles
213	Cannon	Atlanta
33	Levi-Strauss	San Fransisco

Removed “State” in order to reduce redundancies.
City is foreign key to the column City in table State.

Store

<u>Number</u>	City
5	San Fransisco
7	Oakland
8	El Cerrito

Removed “State” in order to reduce redundancies.
City is foreign key to the column City in table State.

State

<u>City</u>	State
San Fransisco	Calif
Los Angeles	Calif
Atlanta	Ga

Created a new table named “State” in order to get rid of the redundancies in tables Supplier and Store.

Employee

Number	First	Last	Salary	Manager	Birthdate	Startdate
157	Tim	Jones	12000	199	1940	1960
1110	Paul	Smith	6000	33	1952	1973
35	Michael	Evans	5000	32	1952	1974

Split “Name” into two columns, “First” and “Last”.

More info stated in the discussion-part of this paper.

Sale was split into 3 tables as it had a couple of FD’s which was not beneficial.

Sold items

Number	Date	Item	Quantity
100581	75-01-15	118	5
100581	75-01-15	120	1
100582	75-01-15	26	1

Number + Item is this tables primary key.

Number is foreign key to the column “Number” in the table “Receipt”.

Cashier

Number	Store	Dept	Employee
100581	7	26	157
100582	8	14	1110
100586	8	53	35

Number is the primary key, and is also the foreign key to Number in the table Sold items.

Receipt

Number	Credit
100581	00000000
100582	24356540
100586	54096831

Number is the primary key, and is also the foreign key to Number in the table Sold items.

Discussion

Employee:

The difference between having one column called “Name” and having 2 columns, “First”- and “Last-name” is that searches for “First name” is somewhat easier with the option of 2 columns. If the database will be used primarily for searching for an employees lastname, then this form is acceptable. Although we opted to make it easier for searches of both first and lastname.

Item:

We assumed that the “Number” column in table “Item” are supposed to be unique, although we found the number 121 to be repeated 2 times.

Sale:

The original table had a couple of problems. The tables contents were: An item/some items had been sold, by an employee at a department, and it was purchased with (presumably) cash or card. So we have 3 different tables:

- Sold items
- Cashier
- Receipt

This way the tables only include 1 of the three contents, which makes it easier to read them.

Discussions has been made with:

- Tobias Lindvall
- Mikael Björkman
- Martin Hjalmarsson
- Mikael Schnell