

Data Normalization with Donut Shop Case

Sourabh Khot | NUID 002754952

Background

A Donut Shop wants create a mobile application for customers to order donuts. As a database designer and developer of a mobile application, I will analyze the data collected by the store in the 'Sales Order' form, and design a normalized logical schema to implement the application database.

First Normal Form (1NF)

After noting all the attributes of the sales order, I excluded 'Subtotal', 'Sales Tax' and 'Total' since these are computed. I got a list of 19 attributes, which I populated with sample data. To achieve 1NF:

1. Atomic value in each column – To achieve this, I did not put all donut items of a single order in one row, but created multiple rows per donut with other information replicated.
2. No repeating groups – No same type of information was repeated in multiple columns. (I assume that home, mobile, and other phone numbers are unique types and hence not in the same group.)
3. Primary key – Analyzing the table, I have created a combination of Order ID and Donut ID as a composite primary key.

Thus, I arrived at the below schema and table with sample data.

1NF Logical Schema:

- ORDER-INFO-1 → {Order ID, Donut ID, Customer ID, Date, First Name, Last Name, Street Address, Apt No., City, State, ZIP, Home Phone, Mobile Phone, Other Phone, Special Handling Notes, Donut Name, Donut Description, Unit Price, Qty}

ORDER-INFO-1:

Composite Primary Key		Customer ID	Date	First Name	Last Name	Street Address	Apt No.	City	State	ZIP	Home Phone	Mobile Phone	Other Phone	Special Handling Notes	Donut Name	Donut Description	Unit Price	Qty
Order ID	Donut ID																	
4532	1	833	May 6, 2014	John	Smith	615 Third St	302	Lilburn	GA	30047	1234567890	9876543210	6789054321	Please include plates and napkins.	Plain	Plain Donut	\$1.50	1
4532	2	833	May 6, 2014	John	Smith	615 Third St	302	Lilburn	GA	30047	1234567890	9876543210	6789054321	Please include plates and napkins.	Glazed	Glazed Donut	\$1.75	5
4532	3	833	May 6, 2014	John	Smith	615 Third St	302	Lilburn	GA	30047	1234567890	9876543210	6789054321	Please include plates and napkins.	Cinnamon	Cinnamon Donut	\$1.75	12
4532	4	833	May 6, 2014	John	Smith	615 Third St	302	Lilburn	GA	30047	1234567890	9876543210	6789054321	Please include plates and napkins.	Chocolate	Chocolate Donut	\$1.75	3
4532	5	833	May 6, 2014	John	Smith	615 Third St	302	Lilburn	GA	30047	1234567890	9876543210	6789054321	Please include plates and napkins.	Sprinkle	Sprinkle Donut	\$1.75	4
4532	6	833	May 6, 2014	John	Smith	615 Third St	302	Lilburn	GA	30047	1234567890	9876543210	6789054321	Please include plates and napkins.	Gluten-Free	Gluten-Free Donut	\$2.00	5

Second Normal Form (2NF)

To derive the second normal form, I had to ensure that all the non-key attributes are functionally dependent on the entire composite primary key, and not parts of it. I found the following functional dependencies on keys:

- {Order ID, Donut ID} → {Qty}
- {Donut ID} → {Donut Name, Donut Description, Unit Price}
- {Order ID} → {Customer ID, Date, First Name, Last Name, Street Address, Apt No., City, State, ZIP, Home Phone, Mobile Phone, Other Phone, Special Handling Notes}

As seen above, all attributes other than 'Qty' are dependent only on parts of the primary key. Hence, I split the data into three tables, removed duplicate rows, assigned a primary key to each table and related foreign keys in other tables, thus meeting 1NF rules as well. The resulting schema and tables are given below.

2NF Logical Schema:

- DONUT-IN-ORDERS → {Order ID, Donut ID, Qty}
- DONUT-INFO → {Donut ID, Donut Name, Donut Description, Unit Price}
- ORDER-INFO-2 → {Order ID, Customer ID, Date, First Name, Last Name, Street Address, Apt No., City, State, ZIP, Home Phone, Mobile Phone, Other Phone, Special Handling Notes}

DONUT-IN-ORDERS:

Composite Primary Key		
Foreign Key	Foreign Key	
<u>Order ID</u>	<u>Donut ID</u>	Qty
4532	1	1
4532	2	5
4532	3	12
4532	4	3
4532	5	4
4532	6	5

DONUT-INFO:

Primary Key			
<u>Donut ID</u>	Donut Name	Donut Description	Unit Price
1	Plain	Plain Donut	\$1.50
2	Glazed	Glazed Donut	\$1.75
3	Cinnamon	Cinnamon Donut	\$1.75
4	Chocolate	Chocolate Donut	\$1.75
5	Sprinkle	Sprinkle Donut	\$1.75
6	Gluten-Free	Gluten-Free Donut	\$2.00

ORDER-INFO-2:

Primary Key	Foreign Key												
<u>Order ID</u>	Customer ID	Date	First Name	Last Name	Street Address	Apt No.	City	State	ZIP	Home Phone	Mobile Phone	Other Phone	Special Handling Notes
4532	833	May 6, 2014	John	Smith	615 Third St	302	Lilburn	GA	30047	1234567890	9876543210	6789054321	Please include plates and napkins.

Third Normal Form (3NF)

As part of the 3NF normalization process, no non-key attribute should be transitively dependent on the primary key, i.e., non-key attributes should not be dependent on each other. In the first two tables of 2NF, all attributes are directly dependent on the primary key.

However, in the ORDERS2 table of 2NF, we can observe that below ten attributes are directly dependent on 'Customer ID', and not on the primary key of 'Order ID': {First Name, Last Name, Street Address, Apt No, City, State, ZIP, Home Phone, Mobile Phone, Other Phone, Special Handling Notes}. There could be multiple 'Order IDs' for the same 'Customer ID', hence we need not repeat the customer details. Therefore, I have separated the order and customer tables and arrived at the below 3NF schema of four tables. These tables correspond to customer info in one table, donut info in another table and order info split into two tables.

Final 3NF Logical Schema:

- DONUT-IN-ORDERS → {Order ID, Donut ID, Qty}
- DONUT-INFO → {Donut ID, Donut Name, Donut Description, Unit Price}
- ORDER-INFO-3 → {Order ID, Customer ID, Date, Special Handling Notes}
- CUSTOMERS-INFO → {Customer ID, First Name, Last Name, Street Address, Apt No., City, State, ZIP, Home Phone, Mobile Phone, Other Phone}

DONUT-IN-ORDERS:

Composite Primary Key		
Foreign Key	Foreign Key	
Order ID	Donut ID	Qty
4532	1	1
4532	2	5
4532	3	12
4532	4	3
4532	5	4
4532	6	5

DONUT-INFO:

Primary Key			
Donut ID	Donut Name	Donut Description	Unit Price
1	Plain	Plain Donut	\$1.50
2	Glazed	Glazed Donut	\$1.75
3	Cinnamon	Cinnamon Donut	\$1.75
4	Chocolate	Chocolate Donut	\$1.75
5	Sprinkle	Sprinkle Donut	\$1.75
6	Gluten-Free	Gluten-Free Donut	\$2.00

ORDER-INFO-3:

Primary Key	Foreign Key		
Order ID	Customer ID	Date	Special Handling Notes
4532	833	May 6, 2014	Please include plates and napkins.

CUSTOMERS-INFO:

Primary Key										
Customer ID	First Name	Last Name	Street Address	Apt No.	City	State	ZIP	Home Phone	Mobile Phone	Other Phone
833	John	Smith	615 Third St	302	Lilburn	GA	30047	1234567890	9876543210	6789054321

All the above tables meet the 1NF, 2NF and 3NF, and hence make up the final logical schema.

<End of Assignment>