

Нормализация базы данных

Нормализация базы данных

Одни и те же данные могут группироваться в таблицы (отношения) различными способами. Группировка атрибутов в отношения должна быть рациональной, т.е. минимизирующей дублирование данных и упрощающей процедуры их обработки и обновления. Устранение избыточности данных является одной из важнейших задач проектирования баз данных и обеспечивается нормализацией.

Нормализация базы данных

Нормализация таблиц (отношений)

– это формальный аппарат ограничений на формирование таблиц (отношений), который позволяет устранить дублирование, обеспечивает непротиворечивость хранимых в базе данных, уменьшает трудозатраты на ведение (ввод, корректировку) базы данных.

1 Нормальная форма (1НФ)

Переменная отношения находится в первой нормальной форме тогда и только тогда, когда в любом допустимом значении отношения каждый его кортеж содержит только одно значение для каждого из атрибутов.

Отношение находится в первой нормальной форме (сокращённо 1НФ), если все его атрибуты атомарны, то есть если ни один из его атрибутов нельзя разделить на более простые атрибуты, которые соответствуют каким-то другим свойствам описываемой сущности.

1 Нормальная форма (1НФ)

- **Первичный ключ.**

Правило: каждая таблица имеет первичный ключ, состоящий из наименьшего возможного количества полей.

Как вы знаете, первичный ключ может состоять из нескольких полей. Вы, к примеру, можете выбрать имя и фамилию в качестве первичного ключа (и надеяться, что эта комбинация будет уникальной всегда). Будет намного более хорошим выбором номер соц. Страхования в качестве первичного ключа, т.к. это единственное поле, которое уникальным образом идентифицирует человека. Еще лучше, когда нет очевидного кандидата на звание первичного ключа, создайте **суррогатный** первичный ключ в виде числового автоинкрементного поля.

1 Нормальная форма (1НФ)

- **Атомарность.**

Правило: поля не имеют дубликатов в каждой записи и каждое поле содержит только одно значение.

1 Нормальная форма (1НФ)

- **Порядок записей не должен иметь значение.**

Правило: порядок записей таблицы не должен иметь значения.

1 Нормальная форма (1НФ)

ФИО	Данные
Иванов Иван Иванович	ПП-119 АСУ Муж. 19.01.1990
Петров Петр Петрович	Э-119 Электронщики Муж. 1991
Васильева Катерина Ильинишна	Прикладная Информатика 1990 Жен ПК 129



Фамилия	Имя	Отчество	Пол	Дата рождения	Группа	Специальность
Иванов	Иван	Иванович	Муж.	19.01.1990	ПП-119	АСУ
Петров	Петр	Петрович	Муж.	20.03.1991	Э-119	Электронщики
Васильева	Катерина	Ильинишна	Жен.	17.04.1990	ПК-129	Прикладная Информатика

2 Нормальная форма (2НФ)

Переменная отношения находится во второй нормальной форме тогда и только тогда, когда она находится в первой нормальной форме и каждый неключевой атрибут неприводимо зависит от ее потенциального ключа.

2 Нормальная форма (2НФ)

- **Избыточность данных.**

Правило: поля с не первичным ключом не должны быть зависимы от первичного ключа.

Фамилия	Имя	Отчество	Пол	Дата рождения	Группа	Специальность
Иванов	Иван	Иванович	Муж.	19.01.1990	ПП-119	АСУ
Петров	Петр	Петрович	Муж.	20.03.1991	Э-119	Электронщики
Васильева	Катерина	Ильинишна	Жен.	17.04.1990	ПК-129	Прикладная Информатика
Петров	Илья	Петрович	Муж.	20.05.1991	Э-119	Электронщики

2 Нормальная форма (2НФ)

Фамилия	Имя	Отчество	Пол	Дата рождения	Группа	Специальность
Иванов	Иван	Иванович	Муж.	19.01.1990	ПП-119	АСУ
Петров	Петр	Петрович	Муж.	20.03.1991	Э-119	Электронщики
Васильев а	Катерин а	Ильинишн а	Жен.	17.04.1990	ПК-129	Прикладная Информатика
Петров	Илья	Петрович	Муж.	20.05.1991	Э-119	Электронщики

Фамилия	Имя	Отчество	Пол	Дата рождения
Иванов	Иван	Иванович	Муж.	19.01.1990
Петров	Петр	Петрович	Муж.	20.03.1991
Васильев а	Катерин а	Ильинишн а	Жен.	17.04.1990
Петров	Илья	Петрович	Муж.	20.05.1991

Специальность
АСУ
Электронщики
Прикладная информатика

Группа
ПП-119
Э-119
ПК-129

2 Нормальная форма (2НФ)

<u>Сотрудник</u>	<u>Должность</u>	<u>Зарплата</u>	<u>Наличие компьютера</u>
Гришин	Кладовщик	20000	Нет
Васильев	Программист	40000	Есть
Иванов	Кладовщик	25000	Нет

<u>Сотрудник</u>	<u>Должность</u>	<u>Зарплата</u>
Гришин	Кладовщик	20000
Васильев	Программист	40000
Иванов	Кладовщик	25000

<u>Должность</u>	<u>Наличие компьютера</u>
Кладовщик	Нет
Программист	Есть

2 Нормальная форма (2НФ)

car_id	brand	type	color	store	price
1	Maserati	Quattroporte	black	Amsterdam South	203000
2	Lada	1118	yellow	Amsterdam South	150
3	Volkswagen	Golf	green	Amsterdam North	14800
4	Volkswagen	Polo	black	Amsterdam West	10200
5	Jaguar	e type	green	The Hague	82399
6	Jaguar	e type	blue	The Hague	22374



car_id	type	color	store	price
1	5	black	1	203000
2	2	yellow	1	150
3	3	green	3	14800
4	4	black	2	10200
5	1	green	4	82399
6	1	blue	4	22374

type_id	brand_id	name
1	3	e type
2	2	1118
3	4	Golf
4	4	Polo
5	1	Quattroporte s

brand_id	name	country_of_origin
1	Maserati	Italy
2	Lada	Russia
3	Jaguar	United Kingdom
4	Volkswagen	Germany

store_id	name	street	hou...	zip...	phone
1	Amsterdam South	Churchil	14	1079HA	020373
2	Amsterdam West	Mercator	27	1056 RT	020838
3	Amsterdam North	Buikslot	76	1031 AB	020387
4	The Hague	Neherstr	82	2491JJ	070387

3 Нормальная форма (3НФ)

Третья нормальная форма связана с **транзитивными зависимостями**. Транзитивные зависимости между полями базы данных существует тогда, когда значения не ключевых полей зависят от значений других не ключевых полей. Чтобы база данных была в третьей нормальной форме, она должна быть во второй нормальной форме.

База данных будет находиться в третьей нормальной форме, если она приведена ко второй нормальной форме и каждый не ключевой столбец независим друг от друга. Если следовать процессу нормализации правильно до этой точки, с приведением к 3НФ может и не возникнуть вопросов. Следует знать, что 3НФ нарушается, если изменив значение в одном столбце, потребуется изменение и в другом столбце.

3 Нормальная форма (3НФ)

Запоминающееся и, по традиции, наглядное резюме определения 3NF Кодда было дано Биллом Кентом: каждый неключевой атрибут «должен предоставлять информацию о ключе, полном ключе и ни о чём, кроме ключа».

Условие зависимости от «полного ключа» неключевых атрибутов обеспечивает то, что отношение находится во второй нормальной форме; а условие зависимости их от «ничего, кроме ключа» — то, что они находятся в третьей нормальной форме.

Крис Дейт говорит о резюме Кента как об «интуитивно привлекательной характеристике» 3NF, и замечает, что с небольшим изменением она может служить и как определение более строгой нормальной формы Бойса-Кодда: «каждый атрибут должен предоставлять информацию о ключе, полном ключе и ни о чём, кроме ключа». Вариант определения 3NF Кента является менее строгим, чем вариант нормальной формы Бойса-Кодда в формулировке Дейта, поскольку первая утверждает только то, что неключевые атрибуты зависят от ключей. Первичные атрибуты (которые являются ключами или их частями) вовсе не должны быть функционально зависимыми; каждый из них предоставляет информацию о ключе предоставлением самого ключа или его части. Здесь следует отметить, что это правило справедливо только для неключевых атрибутов, так как применение его ко всем атрибутам будет полностью запрещать все сложные альтернативные ключи, поскольку каждый элемент такого ключа

3 Нормальная форма (3НФ)

- **Транзитивные зависимости.**

Правило: не может быть транзитивных зависимостей между полями в таблице.

3 Нормальная форма (ЗНФ)

Фамилия	Имя	Отчество	Пол	Дата рождения	Группа	Специальность
Иванов	Иван	Иванович	Муж.	19.01.1990	ПП-119	АСУ
Петров	Петр	Петрович	Муж.	20.03.1991	Э-119	Электронщики
Васильева	Катерина	Ильинишна	Жен.	17.04.1990	ПК-129	Прикладная Информатика
Петров	Илья	Петрович	Муж.	20.05.1991	Э-119	Электронщики

3 Нормальная форма (3НФ)

ID_Student	Фамилия	Имя	Отчество	Пол	Дата рождения	ID_Gruppa
1	Иванов	Иван	Иванович	Муж.	19.01.1990	1
2	Петров	Петр	Петрович	Муж.	20.03.1991	2
3	Васильева	Катерина	Ильинишна	Жен.	17.04.1990	3
4	Петров	Илья	Петрович	Муж.	20.05.1991	2

ID_Gruppa	Группа	ID_Specialnost
1	ПП-119	1
2	Э-119	2
3	ПК-129	3
4	Э-129	2

ID_Specialnost	Специальность
1	АСУ
2	Электронщики
3	Прикладная информатика