Проектирование реляционной базы данных

Разберем принципы реляционной модели и научимся проектировать базу данных.

[Теория реляционных баз данных, нормальные формы](#_t755nhj7uzcu)

[Первая нормальная форма](#_uvs68jqltolt)

[Пример приведения к первой нормальной форме](#_eh2vwxbhc0hf)

[Вторая нормальная форма](#_rxyway75ine5)

[Пример приведения ко второй нормальной форме](#_bb8nzpazxwkt)

[Третья нормальная форма](#_8m9tdmjffmaa)

[Пример приведения к третьей нормальной форме](#_7ksjz1emul0o)

[Нормальная форма Бойса-Кодда](#_2hc4fscyb3tr)

[Четвёртая нормальная форма](#_4y1f5hujyhps)

[Пятая нормальная форма](#_ty6hh778c7om)

[Доменно-ключевая нормальная форма](#_oo0j2lrgwn7a)

[Шестая нормальная форма](#_kncl0atch8ur)

[Практическое применение](#_3h2gm8h7p1py)

[Типы данных в MySQL](#_wmp5ywlhbook)

[Практическое применение](#_6uuthw753cut)

[Ключи](#_uvruapy1kwcw)

[Первичный ключ](#_olm0mm8fumop)

[Уникальный ключ](#_iqpagjvmbnxr)

[Внешний ключ](#_3snogm9ejz6g)

[Индексы и алгоритмы поиска](#_tr9yntmjynt7)

[Практическая работа с учителем](#_1bvq4p7qxw3o)

Практическое [задание](#_g5fwre2xbar7)

[Дополнительные материалы](#_8n4salcam957)

[Используемые ресурсы](#_fh42ram07c5)

# 

# Теория реляционных баз данных, нормальные формы

Принципы реляционной модели были сформулированы в 1969—1970 годах Э. Ф. Коддом (E. F. Codd). Идеи Кодда были впервые публично изложены в статье «A Relational Model of Data for Large Shared Data Banks», ставшей классической.

Обратимся к википедии для определения реляционной модели данных.

Реляционная модель данных (РМД) – логическая модель данных, прикладная теория построения баз данных, которая является приложением к задачам обработки данных таких разделов математики, как теория множеств и логика первого порядка.

С появлением теории реляционных баз данных на смену хранилищам в файловых системах пришли реляционные системы управления базами данных (РСУБД).

Основа теории реляционной модели данных – это нормальные формы (НФ). Для корректного представления и обработки данных существует шесть нормальных форм.

Основные достоинства данного подхода:

1. Позволяет исключить избыточность.
2. Процесс обновления данных становиться простым.
3. Можно легко расширять и дополнять БД, так как она будет отражать сущности реального мира.
4. Проверка целостности данных будет осуществляться за счет БД.

Работу с РСУБД невозможно представить себе без НФ. Давайте рассмотрим шесть НФ.

## Первая нормальная форма

Переменная отношения находится в первой нормальной форме (1НФ), когда в любом допустимом значении отношения каждый его кортеж содержит только одно значение для каждого из атрибутов.

В реляционной модели отношение всегда находится в первой нормальной форме по определению понятия отношение. Что же касается различных таблиц, они могут не быть правильными представлениями отношений и, соответственно, могут не находиться в 1НФ.

### Пример приведения к первой нормальной форме

|  |  |
| --- | --- |
| **Категория** | **Товары** |
| Строительные материалы | Кирпич, цемент, гвозди |
| Отделочные материалы | Гипсокартон, обои, краска |

В ячейке «Товары» данные присутствуют не в атомарном виде, приведем данные к первой нормальной форме.

|  |  |
| --- | --- |
| **Категория** | **Товары** |
| Строительные материалы | Кирпич |
| Строительные материалы | Цемент |
| Строительные материалы | Гвозди |
| Отделочные материалы | Гипсокартон |
| Отделочные материалы | Обои |
| Отделочные материалы | Краска |

## Вторая нормальная форма

Переменная отношения находится во второй нормальной форме, когда она находится в первой нормальной форме, и каждый неключевой атрибут неприводимо (функционально полно) зависит от её потенциального ключа.

### Пример приведения ко второй нормальной форме

|  |  |  |  |
| --- | --- | --- | --- |
| **Категория** | **Дата акции** | **Скидка** | **Товары** |
| Строительные материалы | 21.06.2016 | 10% | Кирпич |
| Строительные материалы | 21.06.2016 | 10% | Цемент |
| Строительные материалы | 21.06.2016 | 10% | Гвозди |
| Отделочные материалы | 25.06.2016 | 20% | Гипсокартон |
| Отделочные материалы | 25.06.2016 | 20% | Обои |
| Отделочные материалы | 25.06.2016 | 20% | Краска |

В этой таблице первичные ключи представлены в виде двух столбцов – категория и дата акции. Схема функционально зависит от категории и не зависит от даты – т. е. зависимость идет от первичного ключа частично. Приведем таблицу ко второму нормальному виду.

|  |  |  |
| --- | --- | --- |
| **Категория** | **Дата акции** | **Скидка** |
| Строительные материалы | 21.06.2016 | 10% |
| Отделочные материалы | 25.06.2016 | 20% |

### 

### 

### 

|  |  |
| --- | --- |
| **Категория** | **Товары** |
| Строительные материалы | Кирпич |
| Строительные материалы | Цемент |
| Строительные материалы | Гвозди |
| Отделочные материалы | Гипсокартон |
| Отделочные материалы | Обои |
| Отделочные материалы | Краска |

## Третья нормальная форма

Переменная отношения находится в третьей нормальной форме тогда и только тогда, когда она находится во второй нормальной форме, и отсутствуют транзитивные функциональные зависимости неключевых атрибутов от ключевых.

### Пример приведения к третьей нормальной форме

Привести к третьей нормальной форме – это значит выделить из таблицы первичный ключ, который будет уникально идентифицировать запись в таблице.

|  |  |
| --- | --- |
| **ID** | **Категория** |
| 1 | Строительные материалы |
| 2 | Отделочные материалы |

|  |  |  |
| --- | --- | --- |
| **Категория** | **Дата акции** | **Скидка** |
| 1 | 21.06.2016 | 10% |
| 2 | 25.06.2016 | 20% |

### 

### 

### 

|  |  |
| --- | --- |
| **Категория** | **Товары** |
| 1 | Кирпич |
| 1 | Цемент |
| 1 | Гвозди |
| 2 | Гипсокартон |
| 2 | Обои |
| 2 | Краска |

## Нормальная форма Бойса-Кодда

Переменная отношения находится в нормальной форме Бойса-Кодда (иначе – в усиленной третьей нормальной форме), когда каждая её нетривиальная и неприводимая слева функциональная зависимость имеет в качестве своего детерминанта некоторый потенциальный ключ.

Другими словами, в таблице должен быть только один первичный ключ.

## Четвёртая нормальная форма

Переменная отношения находится в четвертой нормальной форме, если она находится в нормальной форме Бойса-Кодда и не содержит нетривиальных многозначных зависимостей.

Предположим, что рестораны производят разные виды пиццы, а службы доставки ресторанов работают только в определенных районах города. Составной первичный ключ соответствующей переменной отношения включает три атрибута: {Ресторан, Вид пиццы, Район доставки}.

Такая переменная отношения не соответствует 4НФ, так как существует следующая многозначная зависимость:

{Ресторан} → {Вид пиццы}

{Ресторан} → {Район доставки}

То есть, например, при добавлении нового вида пиццы придется внести по одному новому кортежу для каждого района доставки. Возможна логическая аномалия, при которой определенному виду пиццы будут соответствовать лишь некоторые районы доставки из обслуживаемых рестораном районов.

Для предотвращения аномалии нужно декомпозировать отношение, разместив независимые факты в разных отношениях. В данном примере следует выполнить декомпозицию на {Ресторан, Вид пиццы} и {Ресторан, Район доставки}.

Однако, если к исходной переменной отношения добавить атрибут, функционально зависящий от потенциального ключа, например, цену с учётом стоимости доставки ({Ресторан, Вид пиццы, Район доставки} → Цена), полученное отношение будет находиться в 4НФ и его уже нельзя подвергнуть декомпозиции без потерь.

## Пятая нормальная форма

Переменная отношения находится в пятой нормальной форме (иначе – в проекционно-соединительной нормальной форме), когда каждая нетривиальная зависимость соединения в ней определяется потенциальным ключом (ключами) этого отношения.

Другими словами, если есть сложная связь между зависимыми атрибутами, она должна выноситься в отдельную таблицу. Не нужно делать зависимость трех и более атрибутов. На практике такое встречается очень редко, но придерживаться данного правила необходимо.

## Доменно-ключевая нормальная форма

Переменная отношения находится в ДКНФ, когда каждое наложенное на неё ограничение является логическим следствием ограничений доменов и ограничений ключей, наложенных на данную переменную отношения.

## Шестая нормальная форма

Переменная отношения находится в шестой нормальной форме, когда она удовлетворяет всем нетривиальным зависимостям соединения. Из определения следует, что переменная находится в 6НФ, когда она неприводима, то есть не может быть подвергнута дальнейшей декомпозиции без потерь. Каждая переменная отношения, которая находится в 6НФ, также находится и в 5НФ.

Введена К. Дейтом в его книге как обобщение пятой нормальной формы для хронологической базы данных.

## Практическое применение

Из всех представленных форм практическое применение обычно имеют первые четыре НФ. Наши схемы мы будем создавать на основе четырех НФ.

# 

# 

# Дополнительные материалы

1. <https://habrahabr.ru/post/254773/>

# Используемые ресурсы

Для подготовки данного методического пособия были использованы следующие ресурсы:

1. <https://habrahabr.ru/post/254773/>
2. [https://ru.wikipedia.org/wiki/%D0%A0%D0%B5%D0%BB%D1%8F%D1%86%D0%B8%D0%BE%D0%BD%D0%BD%D0%B0%D1%8F\_%D0%B1%D0%B0%D0%B7%D0%B0\_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85](https://ru.wikipedia.org/wiki/Реляционная_база_данных)
3. [https://ru.wikipedia.org/wiki/%D0%9D%D0%BE%D1%80%D0%BC%D0%B0%D0%BB%D1%8C%D0%BD%D0%B0%D1%8F\_%D1%84%D0%BE%D1%80%D0%BC%D0%B0#.D0.9F.D0.B5.D1.80.D0.B2.D0.B0.D1.8F\_.D0.BD.D0.BE.D1.80.D0.BC.D0.B0.D0.BB.D1.8C.D0.BD.D0.B0.D1.8F\_.D1.84.D0.BE.D1.80.D0.BC.D0.B0\_.281NF.29](https://ru.wikipedia.org/wiki/Нормальная_форма" \l ".D0.9F.D0.B5.D1.80.D0.B2.D0.B0.D1.8F_.D0.BD.D0.BE.D1.80.D0.BC.D0.B0.D0.BB.D1.8C.D0.BD.D0.B0.D1.8F_.D1.84.D0.BE.D1.80.D0.BC.D0.B0_.281NF.29)
4. <http://phpclub.ru/mysql/doc/column-types.html>
5. [http://ru.stackoverflow.com/questions/246659/b-tree-%D0%B8-hash-%D0%B8%D0%BD%D0%B4%D0%B5%D0%BA%D1%81%D1%8B](http://ru.stackoverflow.com/questions/246659/b-tree-и-hash-индексы)