**Первая нормальная форма.**

Нормализация– это процесс (процедура) приведения таблиц базы данных к ряду нормальных форм с целью избежания избыточности в базе данных. Таблицы могут иметь неэффективную или не подходящую структуру, которую нужно нормализовать. Нормализация предусматривает разбивку исходной таблицы (отношения) на несколько новых таблиц (отношений). Так например одно из условий нормализации – атамарность.

Пример:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **id** | **ФИО** | **дата рождения** | **уклон** | **класс** |
| 1 | Петя Петров | 22.12.2010 | математ | 3 |
| 2 | Вася Иванов | 14.08.2010 | математ | 3 |
| 3 | Люся Сидорова | 05.03.2012 | гуманитар | 1 |

Ячейки ФИО и дата рождения содержат не один элемент, преобразуем:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **id** | **Имя** | **Фамилия** | **дд** | **мес** | **год** | **уклон** | **класс** |
| 1 | Петя | Петров | 22 | декабрь | 2010 | математ | 3 |
| 2 | Вася | Иванов | 14 | август | 2010 | мфтемат | 3 |
| 3 | Люся | Сидорова | 5 | май | 2012 | гуманитар | 1 |

Атрибуты должны быть простыми, все используемые домены должны содержать только скалярные значения. Не должно быть повторений строк в таблице.

Правильное применение механизма нормализации к базе данных дает следующие взаимосвязанные преимущества:

* минимизируются затраты на сохранение данных (избыточность данных). Другими словами, уменьшается размер базы данных;
* с возрастанием размера базы данных не так заметно возрастают временные затраты на доступ к данным;
* отсутствуют аномалии модификации в базе данных. К аномалиям модификации относятся аномалии вставки, редактирования и удаления данных.

**Вторая нормальная форма.**

Вторая НФ имеет дело с ключами и зависимостями в схеме. Чтобы база данных находилась во второй НФ, необходимо чтобы ее таблицы удовлетворяли следующим требованиям:

-Таблица должна находиться в первой нормальной форме

-Таблица должна иметь ключ

-Все неключевые столбцы таблицы должны зависеть от полного ключа.

Пример:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **id** | **Имя** | **Фамилия** | **дд** | **мес** | **год** | **уклон** | **класс** |
| 1 | Петя | Петров | 22 | декабрь | 2010 | математ | 3 |
| 2 | Вася | Иванов | 14 | август | 2010 | мфтемат | 3 |
| 3 | Люся | Сидорова | 5 | май | 2012 | гуманитар | 1 |

Уклон программы обучения зависит от класса и года рождения:

|  |  |  |
| --- | --- | --- |
| **год** | **уклон** | **класс** |
| 2010 | математ | 3 |
| 2010 | гуманитар | 3 |
| 2012 | физкульт | 1 |

Чем держать все данные в одной большой таблице, проще представить данные как несколько связанных и логически разделенных отношений. Это проще читать, воспринимать, проектировать и поддерживать.

**Третья нормальная форма.**

3НФ схожа на 2НФ, третья нормальная форма избавляет нас от транзитивных зависимостей: любой столбец таблицы должен зависеть только от первичного ключа (ключевого столбца).

В отношении атрибут фамтлия является первичным ключом:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **id** | **Имя** | **Фамилия** | **дд** | **мес** | **год** | **уклон** | **класс** |
| 1 | Петя | Петров | 22 | декабрь | 2010 | математ | 3 |
| 2 | Вася | Иванов | 14 | август | 2010 | мфтемат | 3 |
| 3 | Люся | Сидорова | 5 | май | 2012 | гуманитар | 1 |

В отношении существуют следующие функциональные зависимости:

Фамилия→год, год→класс, фамилия→класс.

Зависимость Фамилия → классн является транзитивной, следовательно, отношение не находится в 3НФ.  
В результате разделения исходного отношения получаются два отношения, находящиеся в 3НФ:

|  |  |
| --- | --- |
| **год** | **класс** |
| 2010 | 3 |
| 2010 | 3 |
| 2012 | 1 |

|  |  |
| --- | --- |
| **фамилия** | **год** |
| Петя | 2010 |
| Вася | 2010 |
| Люся | 2012 |

**Четверта нормальная форма.**

После того как таблицы базы данных находятся в третьей нормальной форме, мы можем начинать приводить базу данных к четвертой нормальной форме.

Правило четвертой нормальной формы – в таблице не должно быть многозадачгых зависимостей.

Мы должны вынести каждую многозначную зависимость в отдельную таблицу, т.е. разнести независимые друг от друга атрибуты.

Пример: таблица связей ученики, изучение языков, дополнительные занятия:

|  |  |  |
| --- | --- | --- |
| **ученик** | **занятия** | **доп занятия** |
| Петров | англ | плавание |
| Петров | англ | танцы |
| Сидорова | франц | танцы |
| Сидорова | англ | единоборства |
| Иванов | франц | плавание |
| Иванов | франц | футбол |

Зависимость: ученик→занятия

ученик→доп. занятия

Чтобы нормализовать таблицу нужно разбить её на две:

|  |  |
| --- | --- |
| **имя** | **ин яз** |
| Петров | англ |
| Сидорова | франц |
| Сидорова | англ |
| Иванов | франц |

|  |  |
| --- | --- |
| **имя** | **доп занятия** |
| Петров | плавание |
| Петров | танцы |
| Сидорова | танцы |
| Сидорова | единоборства |
| Иванов | плавание |
| Иванов | футбол |

Каждая следующая нормальная форма содержит более строгие правила и критерии. Процес нормализации – это переход от одной нориальной формы к следующей.

Иными словами, процесс перехода от одной нормальной формы к следующей – это усовершенствование базы данных.