**Лабораторная работа №3**

**Нормализация отношений в базе данных**

**Цель работы:** изучить нормальные формы отношений, получить навыки последовательной нормализации отношений.

**Основные теоретические сведения**

Нормализация — процесс преобразования базы данных к виду, отвечающему нормальным формам.

В свою очередь, нормальная форма — совокупность требований, которым должно удовлетворять отношение. Нормальная форма предназначена для устранения из базы избыточных функциональных зависимостей между атрибутами (полями таблиц). Избыточность часто является причиной аномалий, возникших при добавлении, редактировании и удалении строк таблицы. Избыточность устраняется, как правило, за счёт декомпозиции отношений (таблиц), то есть разбиения одной таблицы на несколько.

Типы нормальных форм:

1. первая нормальная форма (1 NF);
2. вторая нормальная форма (2 NF);
3. третья нормальная форма (3 NF);
4. нормальная форма Бойса-Кодда (BCNF);
5. четвёртая нормальная форма (4 NF);
6. пятая нормальная форма (5 NF).

При решении практических задач в большинстве случаев третья нормальная форма является достаточной. Процесс проектирования реляционной базы данных, как правило, заканчивается приведением к 3 NF. 3NF устраняет достаточное количество аномалий, при этом производительность базы данных, а также удобство её использования не снижается, что нельзя сказать обо всех последующих формах.

Рассмотрим подробнее первые три нормальные формы.

Первая нормальная форма (1 NF). Отношение находится в 1 NF тогда и только тогда, когда все входящие в него атрибуты являются атомарными (неделимыми). В таблице такой базы данных не должно быть дублирующих строк, в ячейках должны храниться атомарные значения.

Вторая нормальная форма (2 NF). Отношение находится в 2 NF, если оно находится в 1 NF, а каждый неключевой атрибут функционально полно зависит от первичного ключа.

Таблица такой базы данных должна иметь правильный ключ, по которому можно идентифицировать каждую строку. Если ключ составной, то есть состоит из нескольких столбцов, то все остальные неключевые столбцы должны зависеть от всего ключа. Если какой-то атрибут (столбец) зависит только от части ключа, значит, база данных не находится во второй нормальной форме.

Третья нормальная форма (3 NF).Отношение находится в 3 NF в том и только в том случае, если оно находится в 2 NF, а каждый неключевой атрибут нетранзитивно зависит от первичного ключа.

Транзитивной зависимостью неключевых атрибутов от ключевых называется следующая: A → B и B → C, где A – набор ключевых атрибутов (ключ), B и С – различные множества неключевых атрибутов.

Чтобы нормализовать базу данных до третьей нормальной формы, необходимо сделать так, чтобы в таблицах отсутствовали неключевые столбцы, которые зависят от других неключевых столбцов.

Основные свойства нормальных форм:

1. каждая следующая нормальная форма в некотором смысле является более ограниченной, но лучшей, чем предыдущая;
2. при переходе к следующей нормальной форме положительные свойства предыдущих нормальных свойств сохраняются.

**Задание к работе**

1. Нормализовать структуру базы данных, разработанную в предыдущих лабораторных работах. Доказать соответствие 3 NF, выписав функциональные зависимости для каждой сущности.

2. Привести данное в задании отношение к 3 NF согласно номеру варианта.

**Варианты заданий**

**Вариант 1**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Код сотрудника** | **Номер отдела** | **Отдел** | **Телефон отдела** | **ФИО сотрудника** | **Должность** | **Адрес** |
| 3903 | 1 | Бухгалтерия | 101 | Смирнов А. А. | Бухгалтер | Новая 1-13 |
| 3904 | 2 | Технический отдел | 102 | Егунев Е. А. | Техник | Полевая 34-2 |
| 3905 | 3 | Отдел рекламы | 103 | Ганич И. И. | Менеджер | Майская 2-3 |
| 3906 | 3 | Отдел рекламы | 103 | Авдеева Г. В. | Менеджер | Новая 1-18 |
| 3907 | 1 | Бухгалтерия | 101 | Бобряк Д. Ю. | Главный бухгалтер | Садовая 3-78 |
| 3908 | 4 | Отдел продаж | 104 | Саковец М. О. | Продавец | Лесная 24-2 |
| 3909 | 4 | Отдел продаж | 104 | Ларин М. Р. | Продавец | Полевая 32-6 |
| 3910 | 2 | Технический отдел | 102 | Иванов И. А. | Техник | Лесная 3-56 |
| 3911 | 2 | Технический отдел | 102 | Ковалев Р. А. | Техник | Красная 4-12 |
| 3912 | 4 | Отдел продаж | 104 | Шепелев С. А. | Менеджер | Новая 3-23 |
| 3913 | 4 | Отдел продаж | 104 | Заяц А. А. | Менеджер | Новая 5-12 |

**Вариант 2**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ФИО студента** | **Адрес** | **Код курса** | **Название курса** | **Длительность** | **Стоимость** |
| Смирнов А. А. | Новая 1-13 | 102 | Web-дизайн | 40 | 2600 |
| Егунев Е. А. | Полевая 34-2 | 134 | Компьютерная графика | 56 | 3500 |
| Ганич И. И. | Майская 2-3 | 102 | Web-дизайн | 40 | 2600 |
| Авдеева Г. В. | Новая 1-18 | 156 | HTML | 36 | 1500 |
| Бобряк Д. Ю. | Садовая 3-78 | 107 | MS Access | 24 | 2000 |
| Саковец М. О. | Лесная 24-2 | 134 | Компьютерная графика | 56 | 3500 |
| Ларин М. Р. | Полевая 32-6 | 104 | ООП | 56 | 5000 |
| Иванов И. А. | Лесная 3-56 | 102 | Web-дизайн | 40 | 2600 |
| Ковалев Р. А. | Красная 4-12 | 102 | Web-дизайн | 40 | 2600 |
| Шепелев С. А. | Новая 3-23 | 107 | MS Access | 24 | 2000 |
| Заяц А. А. | Новая 5-12 | 104 | ООП | 56 | 5000 |
| Иванова Е. Е. | Красная 2-3 | 156 | HTML | 36 | 1500 |

**Вариант 3**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Инвентарный номер** | **Название книги** | **Автор** | **Категория** | **ФИО читателя** | **Телефон** | **Дата выдачи** | **Дата возврата** |
| 32 | Кот в сапогах | Ш. Перро | Сказки | Шаров. Д. М. | 12-12-02 | 02.05.13 | 16.05.13 |
| 33 | Windows 95 изнутри | А. Кинг | Пользовательский курс | Шил И. А. | 34-01-23 | 05.05.13 | 08.05.13 |
| 33 | Windows 95 изнутри | А. Кинг | Пользовательский курс | Акимова А. А. | 44-56-11 | 08.05.13 | 22.05.13 |
| 35 | Русалочка | Х. Кристиан Андерсен | Сказки | Левин Н. А. | 12-01-33 | 13.05.13 | 24.05.13 |
| 35 | Русалочка | Х. Кристиан Андерсен | Сказки | Акимова А. А. | 44-56-11 | 30.05.13 |  |
| 36 | Чистый код | К. Мартин Роберт | Программирование | Разуман Н. В. | 71-09-23 | 14.05.13 | 28.06.13 |
| 37 | Совершенный код | М. Стив | Программирование | Шаров Д. М. | 12-12-02 | 28.05.13 | 01.07.13 |
| 38 | 3D Studio MAX в примерах | Э. Потс | Графика | Акимова А. А. | 44-56-11 | 29.05.13 | 12.07.13 |
| 38 | 3D Studio MAX в примерах | Э. Потс | Графика | Черных Л. С. | 34-22-34 | 30.08.13 |  |
| 37 | Совершенный код | М. Стив | Программирование | Шаров Д. М. | 12-12-02 | 31.07.13 |  |

**Вариант 4**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Шифр группы** | **Курс** | **Преподаватель** | **ФИО студента** | **Год рождения** | **Возраст** | **Адрес** |
| 35 | Базовый курс | Дружкова Мария Владимировна | Уварова Мария Анатольевна | 1996 | 17 | Гаражная 16-1 |
| 35 | Базовый курс | Дружкова Мария Владимировна | Щербако Сергей Викторович | 2000 | 13 | Заозерная 11 |
| 35 | Базовый курс | Дружкова Мария Владимировна | Новицкая Любовь Валерьевна | 1998 | 15 | Заозерная 38 |
| 35 | Базовый курс | Дружкова Мария Владимировна | Рахимов Андрей Ришатович | 1997 | 16 | Новая 1-29 |
| 35 | Базовый курс | Дружкова Мария Владимировна | Гнатковский Александр Евгеньевич | 1998 | 15 | Первомайская 12-31 |
| 30 | Программирова- ние | Борисова Мария Леонидовна | Сарайкина Елена Николаевна | 1998 | 15 | Первомайская 16-1 |
| 30 | Программирова- ние | Борисова Мария Леонидовна | Сарайкина Елена Николаевна | 1997 | 16 | Авиационная 8-5 |
| 30 | Программирова- ние | Борисова Мария Леонидовна | Михайленко Артем Олегович | 2000 | 13 | Первомайская 16-4 |
| 34 | Программирова- ние | Борисова Мария Леонидовна | Зайцев Григорий Андреевич | 1999 | 14 | Первомайская 17-41 |
| 34 | Программирова- ние | Борисова Мария Леонидовна | Щенин Леонтий Александрович | 1996 | 17 | Первомайская 15-77 |
| 34 | Программирова- ние | Борисова Мария Леонидовна | Медведева Ирина Ивановна | 2001 | 12 | Первомайская 11-16 |
| 67 | Базовый курс | Токарев Александр Валерьевич | Левин Михаил Андреевич | 2000 | 13 | Первомайская 15-40 |
| 67 | Базовый курс | Токарев Александр Валерьевич | Соловьева Татьяна Александровна | 1998 | 15 | Первомайская 17-50 |
| 67 | Базовый курс | Токарев Александр Валерьевич | Завгарова Юлия Ивановна | 1998 | 15 | Полевая 25-1 |
| 67 | Базовый курс | Токарев Александр Валерьевич | Медведева Марина Александровна | 1999 | 14 | Первомайская 15-86 |
| 67 | Базовый курс | Токарев Александр Валерьевич | Тарабанько Артем Иванович | 1998 | 15 | Новая 1-42 |
| 78 | Подготовитель- ный курс | Дружкова Мария Владимировна | Макарова Татьяна Макаровна | 2000 | 13 | Заозерная 8 |
| 78 | Подготовитель- ный курс | Дружкова Мария Владимировна | Артамонова Алина Сергеевна | 2001 | 12 | Первомайская 17-43 |
| 78 | Подготовитель- ный курс | Дружкова Мария Владимировна | Рогатин Александр Петрович | 2001 | 12 | Первомайская 15-64 |
| 78 | Подготовитель- ный курс | Дружкова Мария Владимировна | Герцун Алексей Иванович | 1999 | 14 | Первомайская 10-18 |

**Вариант 5**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Менеджер** | **Теле-фон** | **Терри-тория** | **Город** | **Клиенты** | **Адрес клиента** | **Наличие долга** | **Оклад** |
| Иванов | 111 | Ленинский р-н | Астрахань | Клиент Л1  Клиент Л2 | Адрес Л1  Адрес Л2 | Нет Нет Нет | 10 |
| Литвинов | 123 | Ленинский р-н | Астрахань | Клиент Л3  Клиент Л4 | Адрес Л3  Адрес Л4 | Да Нет | 10 |
| Сидоров | 101 | Все районы | Ахтубинск | Клиент А1  Клиент А2 | Адрес А1  Адрес А2 | Да Да | 15 |
| Петров | 321 | Кировский р-н | Киров | Клиент К1  Клиент К4 | Адрес К1  Адрес К4 | Нет Нет | 10 |
| Иванова | 110 | Белго-родский р-н | Белгород | Клиент Б2  Клиент Б5 | Адрес Б2  Адрес Б5 | Да Да | 10 |
| Лесная | 345 | Все районы | Москва | Клиент М1  Клиент М2  Клиент М3 | Адрес М1  Адрес М2  Адрес М3 | Нет Нет Да | 20 |
| Большаков | 908 | Все районы | Москва | Клиент М4  Клиент М6  Клиент М12 | Адрес М4  Адрес М6  Адрес М12 | Да Нет Да | 20 |

**Пример выполнения задания**

Пусть дано отношение «R0», имеющее атрибуты «№ сотрудника», «ФИО», «Оклад», «Офис», «Телефон офиса», «Дети», «Имя ребенка», «Возраст ребенка».

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **№ сотрудника** | **ФИО** | **Оклад** | **Офис** | **Телефон офиса** | **Дети** | |
| **Имя** | **Возраст** |
| 211 | Иванов В. К. | 20000 | 12 | 24-24-25 | София Жанна Виктор | 10 7 3 |
| 358 | Петров С. А. | 35000 | 12 | 24-24-25 | Михаил | 5 |
| 360 | Кирова В. М. | 30000 | 5 | 24-25-26 | Анна Владимир | 8 6 |

Табл. 2. Ненормализованное отношение «R0»

Легко заметить, что не все атрибуты отношения «R0» атомарны (неделимы). Очевидно, что у одного сотрудника может быть несколько детей. В отношении дети имеют имя и возраст. Значение этого атрибута для сотрудника под номером 211 содержит несколько кортежей — информацию о трёх детях. Отношение «R0» не находится в 1 NF.

Атрибут «Дети» можно разделить на два более простых атрибута: «Имя ребенка» и «Возраст ребенка». Таким образом, приведем отношение «R0» к 1 NF, вследствие чего получим отношение «R1»:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **№ сотрудника** | **ФИО** | **Оклад** | **Офис** | **Телефон офиса** | **Имя ребенка** | **Возраст ребенка** |
| 211 | Иванов В. К. | 20000 | 12 | 24-24-25 | София | 10 |
| 211 | Иванов В. К. | 20000 | 12 | 24-24-25 | Жанна | 7 |
| 211 | Иванов В. К. | 20000 | 12 | 24-24-25 | Виктор | 3 |
| 358 | Петров С. А. | 35000 | 12 | 24-24-25 | Михаил | 5 |
| 360 | Кирова В. М. | 30000 | 5 | 24-25-26 | Анна | 8 |
| 360 | Кирова В. М. | 30000 | 5 | 24-25-26 | Владимир | 6 |

Табл. 3. Отношение «R1»

|  |  |  |
| --- | --- | --- |
| **№ сотруд-ника** | **Имя ребенка** | **Возраст ребенка** |
| 211 | София | 10 |
| 211 | Жанна | 7 |
| 211 | Виктор | 3 |
| 358 | Михаил | 5 |
| 360 | Анна | 8 |
| 360 | Владимир | 6 |

Ясно, что отношение, находящееся в 1НФ, также может обладать избыточностью. Для её устранения предназначена вторая нормальная форма. Отношение «R1» описывает сразу две сущности — сотрудника и детей сотрудника при первичном ключе «№ сотрудника». Для приведения отношения «R1» ко 2 NF необходимо его разбить на отношения «R2» и «R3»:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **№ сотруд-ника** | **ФИО** | **Оклад** | **Офис** | **Телефон офиса** |
| 211 | Иванов В. К. | 20000 | 12 | 24-24-25 |
| 211 | Иванов В. К. | 20000 | 12 | 24-24-25 |
| 211 | Иванов В. К. | 20000 | 12 | 24-24-25 |
| 358 | Петров С. А. | 35000 | 12 | 24-24-25 |
| 360 | Кирова В. М. | 30000 | 5 | 24-25-26 |
| 360 | Кирова В. М. | 30000 | 5 | 24-25-26 |

Табл. 4. Отношение «R3» Табл. 5. Отношение «R4»

Если при определении первой нормальной формы внимание было нацелено на соблюдение реляционных принципов, а при определении второй нормальной формой в центре внимания был первичный ключ, то для определения третьей нормальной формы внимание должно быть уделено атибутам, которые не являются первичным ключом, то есть неключевым атрибутам. Чтобы нормализовать базу данных до третьей нормальной формы, необходимо сделать так, чтобы в таблицах отсутствовали неключевые атрибуты, которые зависят от других неключевых атрибутов.

Для этого необходимо проверить все неключевые атрибуты. Каждый из них должен зависеть только от первичного ключа.

Посмотрим еще раз на отношение «R3». Можно увидеть, что атрибут «Телефон офиса» не зависит напрямую от первичного ключа. Мы это выяснили, когда задали себе вопрос: «Каким образом телефон офиса связан с сотрудником?» Ответ: «Телефон офиса имеется у того офиса, в котором работает сотрудник».

Отсюда следует, что атрибут «Телефон офиса» не связан на прямую с сотрудником, он связан напрямую со с атрибутом «Офис», который напрямую связан с сотрудником, ведь сотрудник работает в каком-то конкретном офсе. Это и есть транзитивная зависимость, когда один неключевой атрибут связан с первичным ключом через другой неключевой атрибут.

Для приведения отношения к 3 NF нужно отношение «R3» разбить на отношения «R4» и «R5»:

|  |  |  |  |
| --- | --- | --- | --- |
| **№ сотрудника** | **ФИО** | **Оклад** | **Офис** |
| 211 | Иванов В. К. | 20000 | 12 |
| 358 | Петров С. А. | 35000 | 12 |
| 360 | Кирова В. М. | 30000 | 5 |

|  |  |
| --- | --- |
| **Офис** | **Телефон офиса** |
| 12 | 24-24-25 |
| 5 | 24-25-26 |

Табл. 5. Отношение «R4» Табл. 6. Отношение «R5»

В процессе выполнения лабораторной работы были изучены нормальные формы отношений и их особенности, а также получены навыки последовательной нормализации отношений (первая, вторая и третья нормальные формы).