

Проектирование баз данных



Основные понятия



База данных

это организованная структура, предназначенная для хранения информации

Реляционная модель

модель, которая описывает как организовать данные в таблицах и как определить связи между этими таблицами

Примеры РСУБД

Oracle, MySQL, DB2, Microsoft SQL server

CSV VS Таблицы

```
1,Алиева,Аня
2,Бирилло,Саша
3,Гузлов,Лёша
4,Кривицкая,Полина
5,Лебедева,Вика
6,Микулич,Егор
7,Рубацкий,Андрей
8,Соколенко,Дима
9,Супрунович,Вероника
10,Хацкевич,Максим
11,Чапурин,Ваня
12,Ярмолук,Паша
```

id	last_name	first_name
1	Алиева	Аня
2	Бирилло	Саша
3	Гузлов	Лёша
4	Кривицкая	Полина
5	Лебедева	Вика
6	Микулич	Егор
7	Рубацкий	Андрей
8	Соколенко	Дима
9	Супрунович	Вероника
10	Хацкевич	Максим
11	Чапурин	Ваня
12	Ярмолук	Паша



01

Использование ключей

02

Отсутствие избыточности данных

03

Ограничение ввода

Характеристики реляционных баз данных



04

Поддержание целостности данных

05

Назначение прав

06

Переносимость

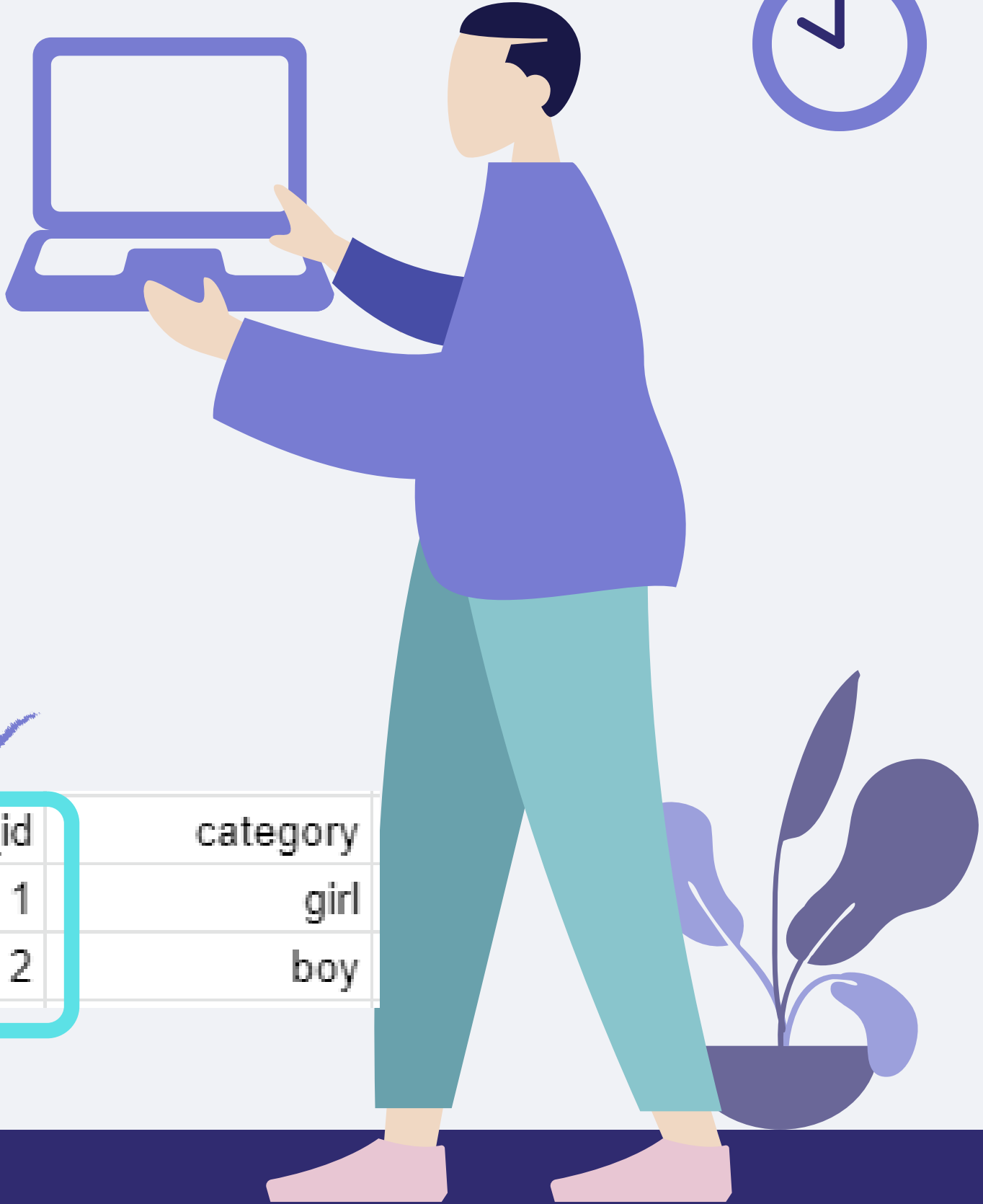
Характеристики реляционных баз данных



Использование ключей

id	last_name	first_name	user_group
1	Алиева	Аня	1
2	Бирилло	Саша	2
3	Гузлов	Лёша	2
4	Кривицкая	Полина	1
5	Лебедева	Вика	1
6	Микулич	Егор	2
7	Рубацкий	Андрей	2
8	Соколенко	Дима	2
9	Супрунович	Вероника	1
10	Хацкевич	Максим	2
11	Напурин	Ваня	2
12	Ярмолюк	Паша	2

group_id	category
1	girl
2	boy

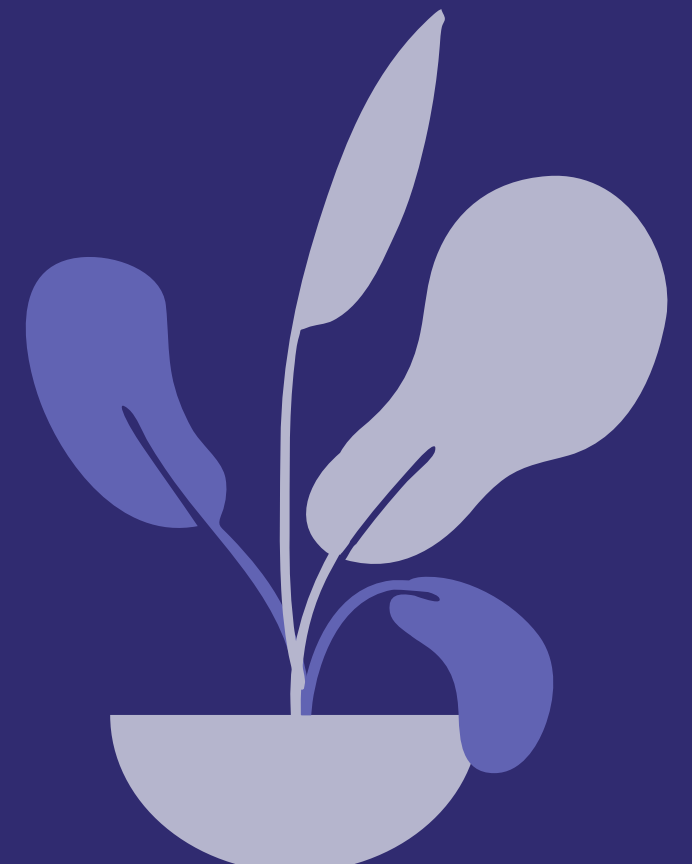


Первичный ключ

- поле или набор полей, однозначно (уникально) идентифицирующих запись. Первичный ключ должен быть минимально достаточным: в нем не должно быть полей, удаление которых из первичного ключа не отразится на его уникальности.

Внешний ключ

- столбец или комбинация столбцов, значения которых соответствуют Первичному ключу в другой таблице





- Номер заказа, который вы получаете при покупке в интернет-магазине может быть первичным ключом какой-нибудь таблицы заказов в базе данных этого магазина, т.к. он является уникальным значением.
- Номер социального страхования может быть первичным ключом в какой-нибудь таблице в базе данных государственного учреждения, т.к. она также как и в предыдущем примере уникален.
- Числовой номер клиента часто используется как первичный ключ в таблице клиентов.

Примеры

Связывание таблиц с помощью внешних ключей

- один — один (1:1);
- один — много (1:M);
- много — один (M: 1);
- много — много (M:M или M:N).

Один-ко-многим

имеет место в случае, когда одной записи основной таблицы соответствует несколько записей вспомогательной таблицы

Таблица клиентов:

- **customer_id** (primary key) – идентификатор клиента
- **first_name** — имя
- **last_name** — отчество
- **address** — адрес
- **username** – регистрационное имя пользователя (логин)
- **password** – пароль

Таблица заказов:

- **order_id** (primary key) – идентификатор заказа
- **order_date** – дата и время заказа
- **customer** – клиент, который сделал заказ

Как распознать связь один ко многим?



Если у вас есть две
сущности спросите себя:

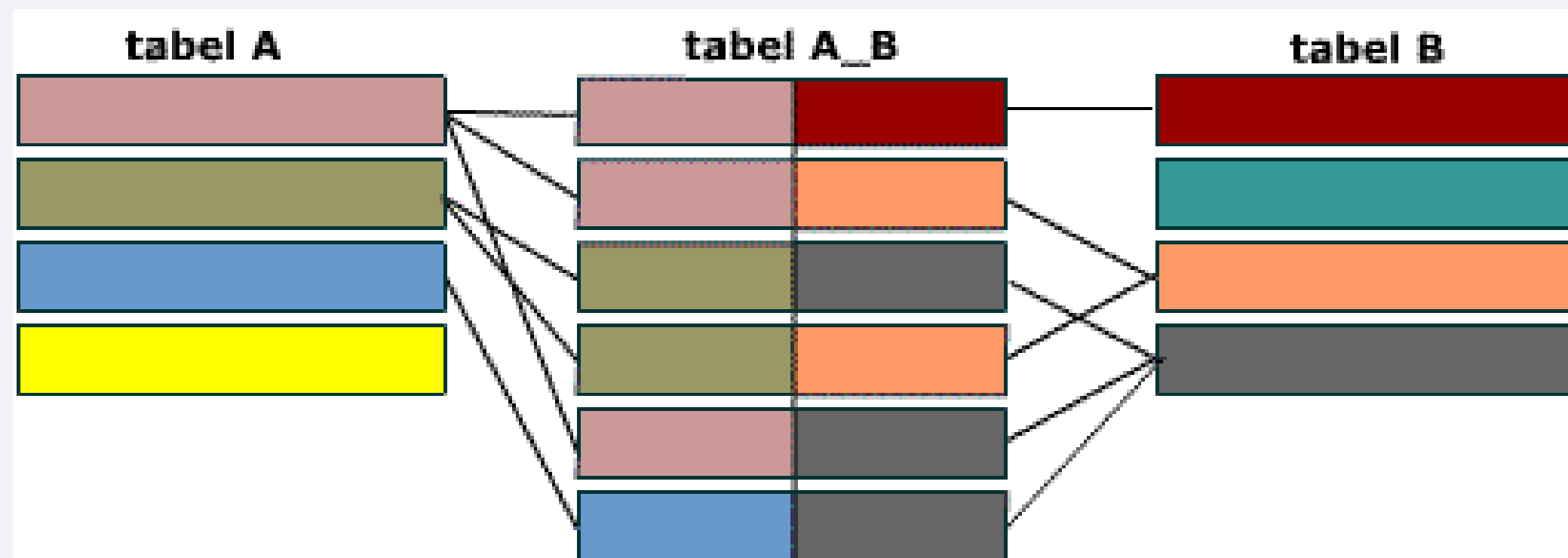
- Сколько объектов из В могут относиться к объекту А?
- Сколько объектов из А могут относиться к объекту из В?

Если на первый вопрос ответ –
множество, а на второй – один (или
возможно, ни одного), то вы имеете
дело со связью один-ко-многим.

Многие-ко-многим

это связь, при которой множественным записям из одной таблицы (A) могут соответствовать множественные записи из другой (B)

Связь многие-ко-многим создается с помощью трех таблиц. Две таблицы – “источника” и одна соединительная таблица. Первичный ключ соединительной таблицы A_B – составной. Она состоит из двух полей, двух внешних ключей, которые ссылаются на первичные ключи таблиц A и B.





Нормализация баз данных

процесс удаления избыточных данных.

метод проектирования базы данных, который
позволяет привести базу данных к минимальной
избыточности



Зачем нормализовать базу данных?



Устранение
аномалий

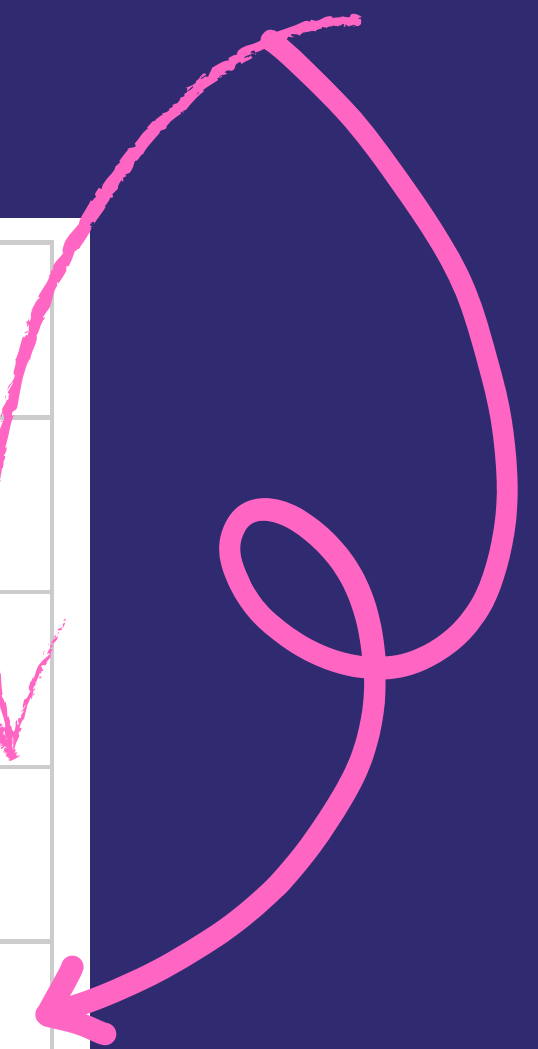
Повышение
производительности

Повышение удобства
управления данными

Идентификатор предмета	Наименование предмета	Материал
1	Стул	Металл
2	Стол	Массив дерева
3	Кровать	ЛДСП
4	Шкаф	Массив дерева
5	Комод	ЛДСП

Аномалия

Идентификатор предмета	Наименование предмета	Материал
1	Стул	Металл
2	Стол	Натуральное дерево
3	Кровать	ЛДСП
4	Шкаф	Массив дерева
5	Комод	ЛДСП



Декомпозиция

Идентификатор предмета	Наименование предмета	Идентификатор материала
1	Стул	2
2	Стол	1
3	Кровать	3
4	Шкаф	1
5	Комод	3



Solution

Идентификатор материала	Материал
1	Массив дерева
2	Металл
3	лдсп

Первая нормальная форма (1NF)

основное правило - атомарность значений

- В таблице не должно быть дублирующих строк
- В столбце хранятся данные одного типа
- Отсутствуют массивы и списки в любом виде

Таблица сотрудников в ненормализованном виде

Сотрудник	Контакт
Иванов И.И.	123-456-789, 987-654-321
Сергеев С.С.	Рабочий телефон 555-666-777, Домашний телефон 777-888-999
John Smith	123-456-789
John Smith	123-456-789

Декомпозиция

Сотрудник	Телефон	Тип телефона
Иванов И.И.	123-456-789	
Иванов И.И.	987-654-321	
Сергеев С.С.	555-666-777	Рабочий телефон
Сергеев С.С.	777-888-999	Домашний телефон
John Smith	123-456-789	



Solution

Требования Второй нормальной формы (2NF)



- Таблица должна находиться в первой нормальной форме
- Таблица должна иметь ключ
- Все неключевые столбцы таблицы должны зависеть от полного ключа (в случае если он составной)

Таблица сотрудников в первой нормальной форме

ФИО	Должность	Подразделение	Описание подразделения
Иванов И.И.	Программист	Отдел разработки	Разработка и сопровождение приложений и сайтов
Сергеев С.С.	Бухгалтер	Бухгалтерия	Ведение бухгалтерского и налогового учета финансово- хозяйственной деятельности
John Smith	Продавец	Отдел реализации	Организация сбыта продукции

Декомпозиция

Идентификатор проекта	Название проекта	Срок проекта (мес.)
1	Внедрение приложения	8
2	Открытие нового магазина	12

Идентификатор участника	Участник	Должность
1	Иванов И.И.	Программист
2	Сергеев С.С.	Бухгалтер
3	John Smith	Менеджер



Solution

Идентификатор проекта	Идентификатор участника
1	1
1	2
1	3
2	2
2	3

Третья нормальная форма(3NF)

2NF + отсутствует зависимость неключевых полей от других неключевых полей

Табельный номер	ФИО	Должность	Подразделение	Описание подразделения
1	Иванов И.И.	Программист	Отдел разработки	Разработка и сопровождение приложений и сайтов
2	Сергеев С.С.	Бухгалтер	Бухгалтерия	Ведение бухгалтерского и налогового учета финансово-хозяйственной деятельности
3	John Smith	Продавец	Отдел реализации	Организация сбыта продукции

Декомпозиция

Табельный номер	ФИО	Должность	Подразделение
1	Иванов И.И.	Программист	1
2	Сергеев С.С.	Бухгалтер	2
3	John Smith	Продавец	3

Идентификатор подразделения	Подразделение	Описание подразделения
1	Отдел разработки	Разработка и сопровождение приложений и сайтов
2	Бухгалтерия	Ведение бухгалтерского и налогового учета финансово-хозяйственной деятельности
3	Отдел реализации	Организация сбыта продукции



Solution

Четвертая нормальная форма (4NF)

- требование четвертой нормальной формы (4NF) заключается в том, чтобы в таблицах отсутствовали нетривиальные многозначные зависимости.

Идентификатор курса	Название курса
1	SQL
2	Python
3	JavaScript

Идентификатор преподавателя	ФИО
1	Иванов И.И.
2	Сергеев С.С.
3	John Smith

Идентификатор аудитории	Название аудитории
1	101
2	203
3	305
4	407
5	502

Курс	Преподаватель	Аудитория
SQL	Иванов И.И.	101
SQL	Иванов И.И.	203
SQL	Сергеев С.С.	305
SQL	Сергеев С.С.	407
Python	John Smith	502
Python	John Smith	305

Декомпозиция

Курс	Преподаватель
SQL	Иванов И.И.
SQL	Сергеев С.С.
Python	John Smith

Курс	Аудитория
SQL	101
SQL	203
SQL	305
SQL	407
Python	502
Python	305



Solution

Студент	Курс	Хобби
Иванов И.И.	SQL	Футбол
Иванов И.И.	Java	Хоккей
Сергеев С.С.	SQL	Волейбол
Сергеев С.С.	SQL	Теннис
John Smith	Python	Футбол
John Smith	Java	Теннис