

Лекция 5.

Нормализация реляционной базы данных.

Нормализация - формальная процедура, в ходе которой одна таблица разбивается на две или несколько по определенным правилам.

Нормализация - формальная процедура, в ходе которой одна таблица разбивается на две или несколько по определенным правилам.

Задачи нормализации:

- Исключение из таблиц повторяющейся информации.
- Создание структуры, в которой предусмотрена возможность ее будущих изменений и влияние таких изменений на приложения минимально.

Дублирование и избыточность данных

Книга	Студент
Книга 1	Иванов
Книга 2	Петров
Книга 3	Иванов

Данные **дублированные**, но не избыточные

Дублирование и избыточность данных

Книга	Студент
Книга 1	Иванов
Книга 2	Петров
Книга 3	Иванов

Данные **дублированные**, но не избыточные

Книга	Студент	Телефон
Книга 1	Иванов	9631440779
Книга 2	Петров	9170550011
Книга 3	Иванов	9631440779

Данные по телефонам
избыточные

Дублирование и избыточность данных

Книга	Студент
Книга 1	Иванов
Книга 2	Петров
Книга 3	Иванов

Данные **дублированные**, но не избыточные

Книга	Студент	Телефон
Книга 1	Иванов	9631440779
Книга 2	Петров	9170550011
Книга 3	Иванов	9631440779

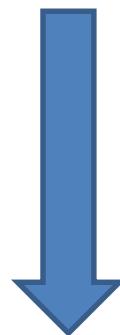
Данные по телефонам
избыточные

В чем **опасность избыточности данных** в БД?

Дублирование и избыточность данных

Книга	Студент	Телефон
Книга 1	Иванов	9631440779
Книга 2	Петров	9170550011
Книга 3	Иванов	9631440779

Данные по телефонам
избыточные



Разбиваем на две таблицы:

Книга	Студент
Книга 1	Иванов
Книга 2	Петров
Книга 3	Иванов

Студент	Телефон
Иванов	9631440779
Петров	9170550011

Нормальные формы

Первая НФ: атомарные атрибуты. Требует, чтобы каждое поле таблицы:

- было неделимым (не должно делиться на более мелкие значения)
- не содержало повторяющихся групп.

Нормальные формы

Первая НФ: атомарные атрибуты. Требуется, чтобы каждое поле таблицы:

- было неделимым (не должно делиться на более мелкие значения)
- не содержало повторяющихся групп.

Вторая НФ: отсутствие зависимостей частичного ключа. Требуется, чтобы все поля таблицы зависели от первичного ключа, т.е. чтобы первичный ключ однозначно определял запись и не был избыточным.

Те поля, которые зависят только от части первичного ключа, должны быть выделены в отдельные таблицы.

Нормальные формы

Третья НФ: **устранение транзитивных зависимостей.**

Требует, чтобы в таблице не имелось транзитивных зависимостей между неключевыми полями, т.е. чтобы значение любого поля таблицы, не входящего в первичный ключ, не зависело от значения другого поля, не входящего в первичный ключ.

Первая НФ

Студент	Телефон
Иванов	9631440779; 9780231200
Петров	9170550011; 9051231234

Первая НФ

Студент	Телефон
Иванов	9631440779; 9780231200
Петров	9170550011; 9051231234

Есть повторяющиеся группы –
нарушение 1 НФ

Первая НФ

Студент	Телефон
Иванов	9631440779; 9780231200
Петров	9170550011; 9051231234

Есть повторяющиеся группы –
нарушение 1 НФ

Студент	Телефон
Иванов	9631440779
Иванов	9780231200
Петров	9051231234
Петров	9170550011

Таблица приведена к 1 НФ

Пример

Спроектировать БД для автоматизации отпуска товаров со склада

Накладная № 5

Дата: **03.12.2016** Покупатель: **ООО "Друг"** Адрес: **г. Саранск**

Наименование	Кол-во	Ед. изм.	Цена	Сумма
Тушенка	1000	Банка	50	50000
Сахар	100	Кг	30	3000

Итого: **53000**

Собираем все в одну таблицу «Отпуск товара»

Номер	Дата	Покупатель	Адрес	Наименование	Ед_изм	Цена	Кол-во	Сумма
5	03.12.2016	ООО Друг	Саранск	Тушенка	Банка	50	1000	50000
5	03.12.2016	ООО Друг	Саранск	Сахар	Кг	30	100	3000



Отпуск_товаров
Номер
Дата
Покупатель
Адрес
Наименование
Ед_изм
Цена
Количество
Сумма

Собираем все в одну таблицу «Отпуск товара»

Номер	Дата	Покупатель	Адрес	Наименование	Ед_изм	Цена	Кол-во	Сумма
5	03.12.2016	ООО Друг	Саранск	Тушенка	Банка	50	1000	50000
5	03.12.2016	ООО Друг	Саранск	Сахар	Кг	30	100	3000



Отпуск_товаров
Номер
Дата
Покупатель
Адрес
Наименование
Ед_изм
Цена
Количество
Сумма

Начинаем нормализацию:
Таблица приведена к 1 НФ ?

Собираем все в одну таблицу «Отпуск товара»

Номер	Дата	Покупатель	Адрес	Наименование	Ед_изм	Цена	Кол-во	Сумма
5	03.12.2016	ООО Друг	Саранск	Тушенка	Банка	50	1000	50000
5	03.12.2016	ООО Друг	Саранск	Сахар	Кг	30	100	3000



Отпуск_товаров
Номер
Дата
Покупатель
Адрес
Наименование
Ед_изм
Цена
Количество
Сумма

Начинаем нормализацию:
Таблица приведена к 1 НФ ? **Да**

Первичный ключ?

Выделение таблицы «Товары»

Отпуск_товаров
Номер
Наименование
Дата
Покупатель
Адрес
Ед_изм
Цена
Количество
Сумма

← Первичный ключ

Какие поля зависят только от части
первичного ключа?

Выделение таблицы «Товары»

Отпуск_товаров
Номер
Наименование
Дата
Покупатель
Адрес
Ед_изм
Цена
Количество
Сумма

← Первичный ключ

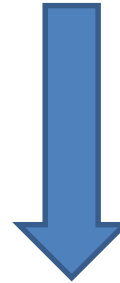
«Ед_изм» и «Цена» зависят только от «Наименование»

Выделение таблицы «Товары»

Отпуск_товаров
Номер
Наименование
Дата
Покупатель
Адрес
Ед_изм
Цена
Количество
Сумма

← Первичный ключ

«Ед_изм» и «Цена» зависят только от «Наименование»

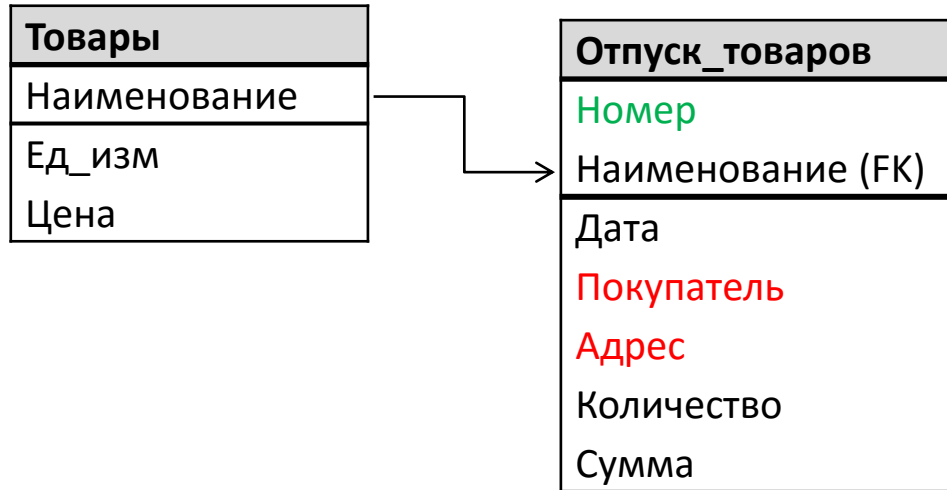


Товары
Наименование
Ед_изм
Цена

Отпуск_товаров
Номер
Наименование (FK)
Дата
Покупатель
Адрес
Количество
Сумма

Выделение таблицы «Покупатели»

«Покупатель» и «Адрес» зависят только от номера накладной

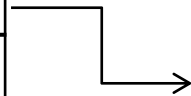


Выделение таблицы «Покупатели»

«Покупатель» и «Адрес» зависят только от номера накладной



Товары
Наименование
Ед_изм
Цена



Отпуск_товаров
Номер
Наименование (FK)
Дата
Количество
Сумма

Покупатели
Покупатель
Адрес

Выделение таблицы «Накладные»

«Дата» зависит только от номера накладной

Товары
Наименование
Ед_изм
Цена

Отпуск_товаров
Номер
Наименование (FK)
Дата
Количество
Сумма

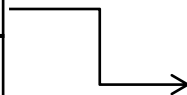
Покупатели
Покупатель
Адрес

Выделение таблицы «Накладные»

«Дата» зависит только от номера накладной



Товары
Наименование
Ед_изм
Цена



Отпуск_товаров
Номер
Наименование (FK)
Количество
Сумма

Покупатели
Покупатель
Адрес

Накладные
Номер
Дата

Устанавливаем связи



Приведена ли БД к 2-й НФ?

Устанавливаем связи



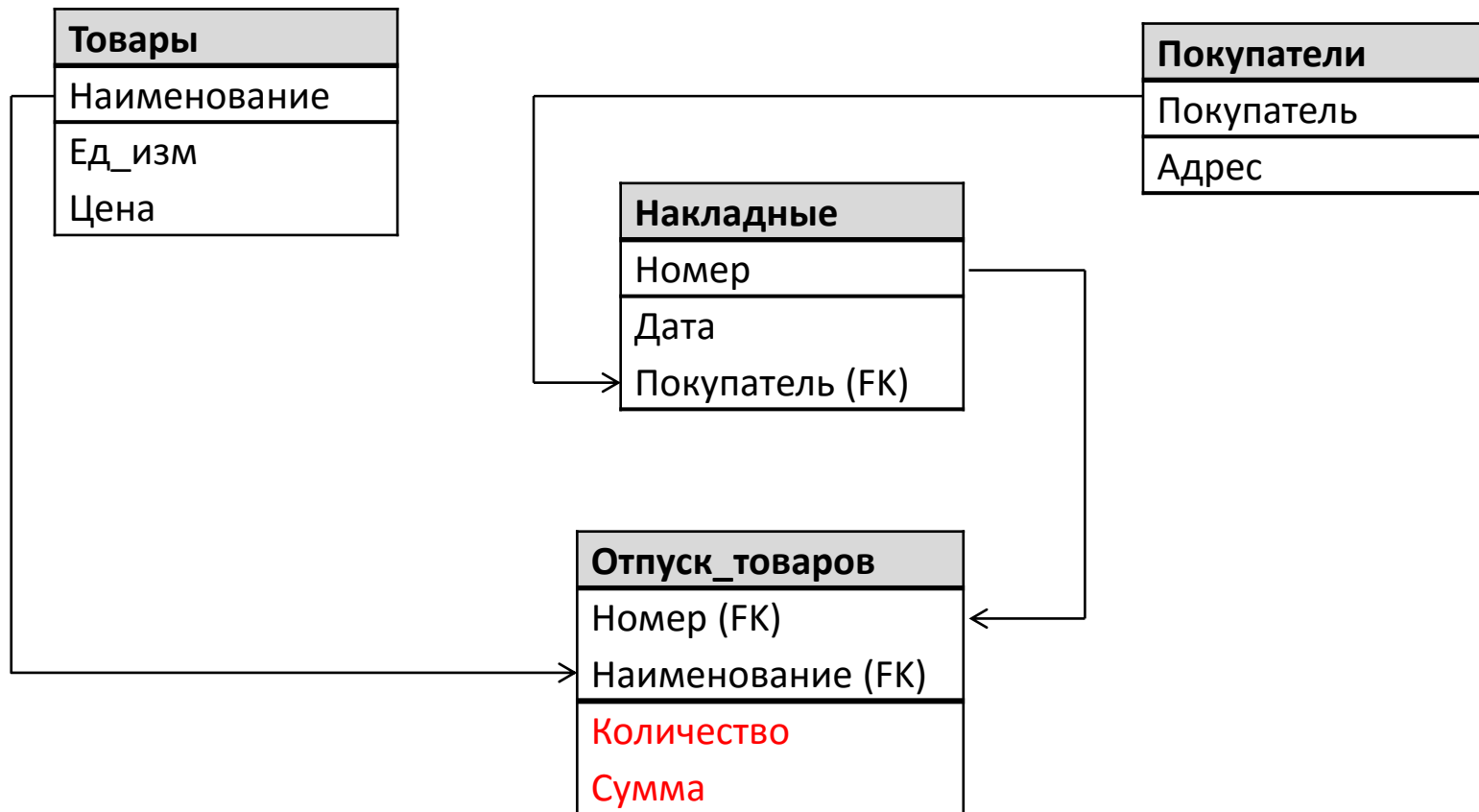
БД приведена к 2-й НФ

Устанавливаем связи



Приведена ли БД к 3-й НФ?

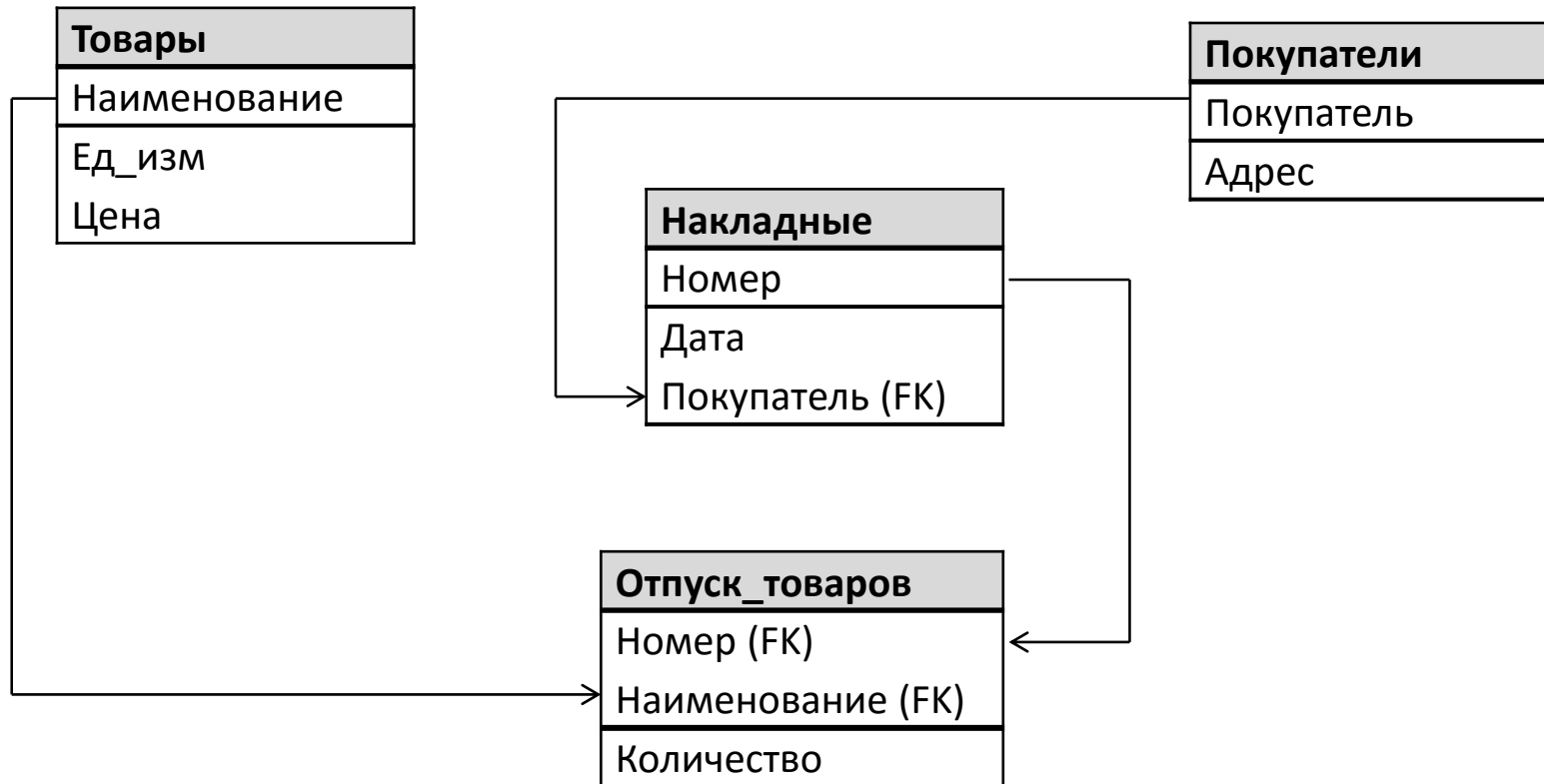
Устанавливаем связи



«Сумма» зависит от «количества»

Удаляем лишние поля

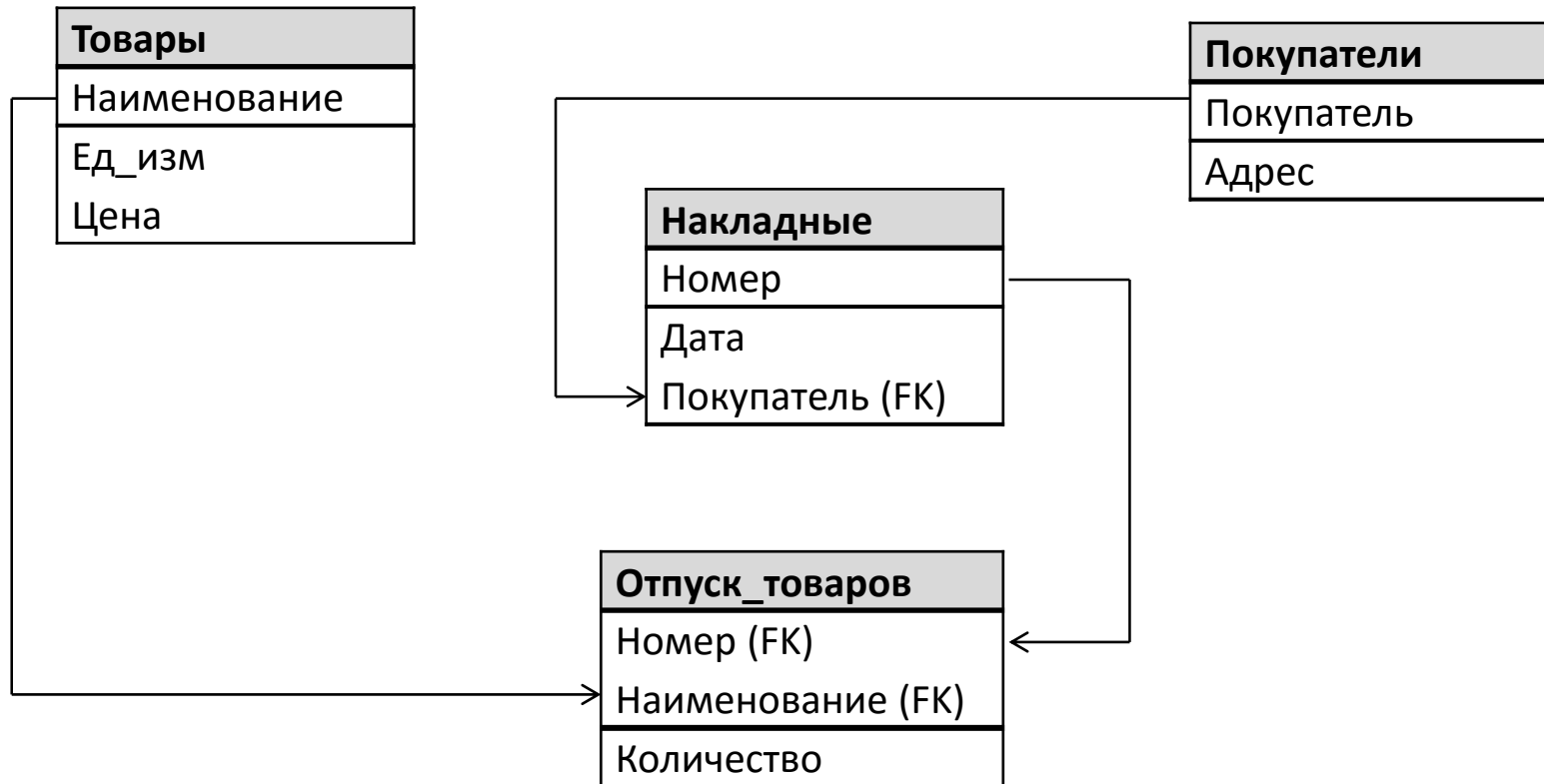
"Отпуск_товаров"."Сумма"="Отпуск_товаров"."Количество"*"Товары"."Цена"



БД приведена к 3-й НФ

Удаляем лишние поля

"Отпуск_товаров"."Сумма"="Отпуск_товаров"."Количество"*"Товары"."Цена"



БД приведена к 3-й НФ

Результаты нормализации

Плюс. Таблицы содержат только один элемент избыточных данных – это поле связи, присутствующее одновременно у родительской и дочерних таблиц.

Минус. Сильно увеличивается число таблиц:

- Теряется восприятие системы в целом.
- Усложняются запросы, приходится объединять данные из разных таблиц (скорость выполнения падает)