



ОНЛАЙН-ОБРАЗОВАНИЕ


Онлайн-образование

Не забыть включить запись!





Меня хорошо видно && слышно?

Ставьте  , если все хорошо
Напишите в чат, если есть проблемы

Проектирование БД

курс “SQL Server Developer”



Коробков Виктор

Правила вебинара



Активно участвуем



Задаем вопрос в чат или голосом



Off-topic обсуждаем в Slack



Вопросы вижу в чате, могу ответить не сразу

Цели вебинара | После занятия вы сможете

- 1 Перечислить этапы проектирования баз данных
- 2 Выполнять нормализацию таблиц
- 3 Производить ER-моделирование базы данных

Смысл | зачем вам это уметь

Для создания реального рабочего проекта базы данных в любой предметной области

Маршрут вебинара

1 Немного о проекте

2 Реляционная модель данных

3 Этапы проектирования баз данных

4 Нормализация таблиц

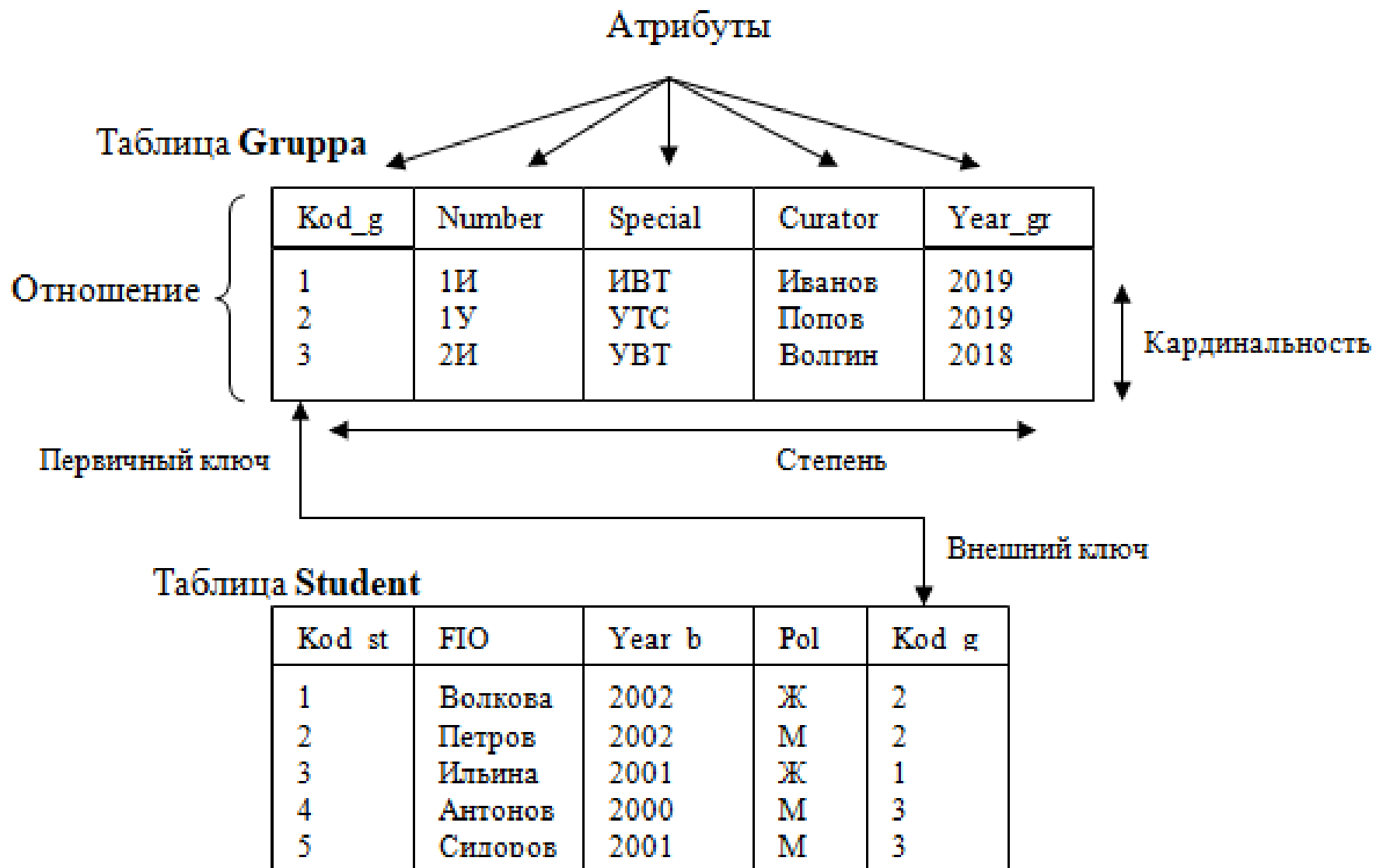
Описание проекта

1. Проект БД по вашей идее: свой проект или кусочек с работы, который вы хотите пересмотреть.
2. Порядка 10-15 таблиц, можно меньше или больше, это не жестко заданное требование.
3. Желательно использовать SQL Server 2017, если это не прод проект с жестко заданной версией БД.
4. Делаем структуру БД, с описанием, схемой.
5. В конце модуля небольшая презентация, с объяснением интересных частей проекта и обсуждением принятых решений и вариантов.

Описание проекта

Коллеги, вопросы ?

Реляционная модель данных



Основные понятия:

отношение;

атрибут;

тип данных;

домен;

кортеж;

первичный ключ;

внешний ключ.

Реляционная модель данных

Свойства отношений:

1. Отношение имеет имя, которое отличается от имен всех других отношений в реляционной схеме.
2. Каждая ячейка отношения содержит только одно элементарное (неделимое/атомарное) значение.
3. Каждый атрибут имеет уникальное имя в пределах отношения.
4. Значения атрибута берутся из одного и того же домена;
5. Каждый кортеж является уникальным, т.е. дубликатов кортежей быть не может.
6. Порядок следования атрибутов не имеет значения.
7. Порядок следования кортежей в отношении не имеет значения.

Реляционная модель данных

Суперключ (superkey) – это атрибут или множество атрибутов, которое единственным образом идентифицирует каждый кортеж данного отношения.

!!! Суперключ может содержать дополнительные атрибуты, которые необязательны для уникальной идентификации кортежа. **!!!**

Реляционная модель данных

Потенциальный ключ (candidate key) – это суперключ, который не содержит подмножества, также являющегося суперключом данного отношения.

Потенциальный ключ K для данного отношения R обладает двумя свойствами:

- уникальность – в каждом кортеже отношения R значение ключа K единственным образом идентифицируют этот кортеж;
- неприводимость – никакое допустимое подмножество ключа K не обладает свойством уникальности.

!!! Отношение может иметь несколько потенциальных ключей. **!!!**

Реляционная модель данных

Первичный ключ (primary key) – это потенциальный ключ, который выбран для уникальной идентификации кортежей внутри отношения.

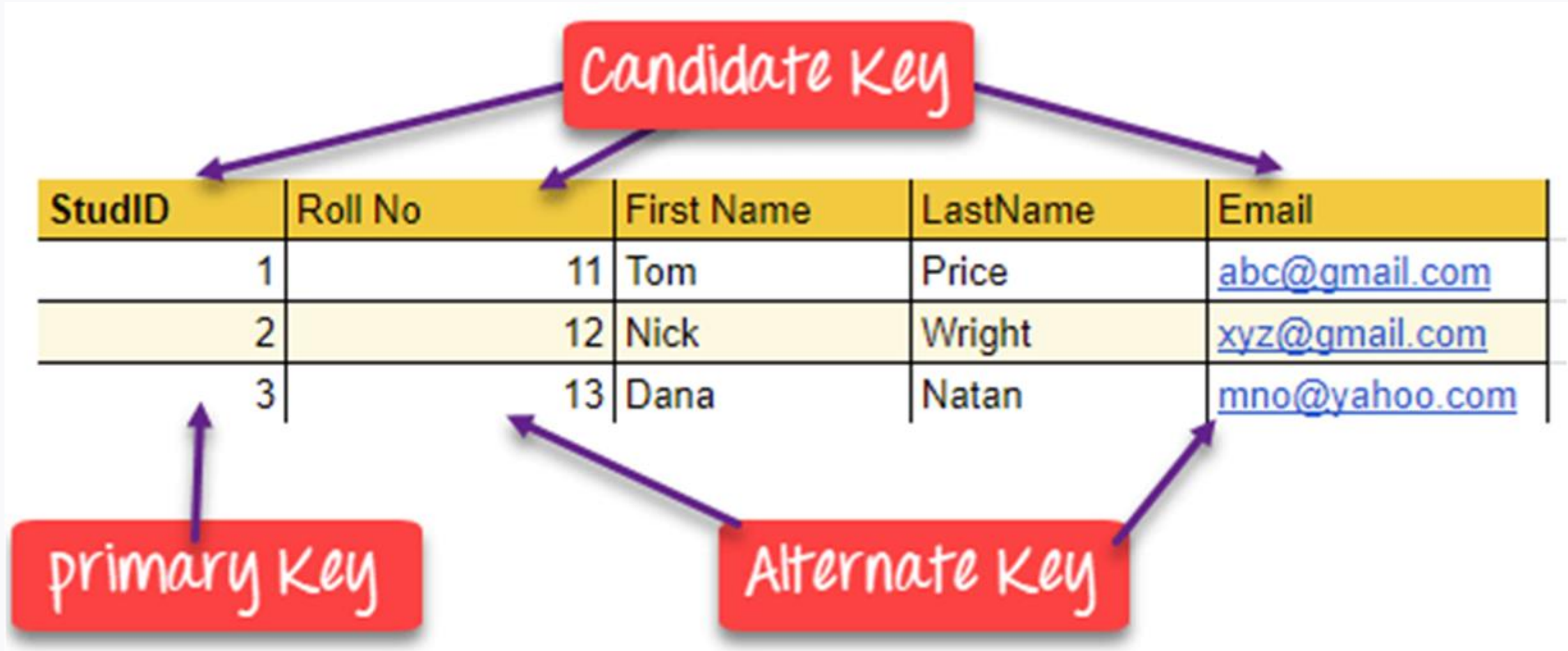
Альтернативный ключ (alternate key) – это потенциальный ключ, который не выбран в качестве первичного ключа.

Внешний ключ (foreign key) – это атрибут или множество атрибутов внутри отношения, которое соответствует потенциальному ключу некоторого (может быть, того же самого) отношения.

Требования целостности данных

1. **Целостность сущностей** – в базовом отношении ни один атрибут первичного ключа не может содержать отсутствующих значений, обозначаемых как NULL.
2. **Ссылочная целостность** – если в отношении существует внешний ключ, то значение внешнего ключа должно либо соответствовать значению потенциального ключа некоторого кортежа в его базовом отношении либо внешний ключ должен полностью состоять из значений NULL.
3. **Корпоративные ограничения целостности** – дополнительные правила поддержки целостности данных, определяемые пользователями или администраторами базы данных.
4. Соблюдения правила при **удалении** данных:
 - запрет удаления данных;
 - изменение значения внешнего ключа на NULL;
 - каскадное удаление (удалить ссылающиеся записи).

Пример ключей

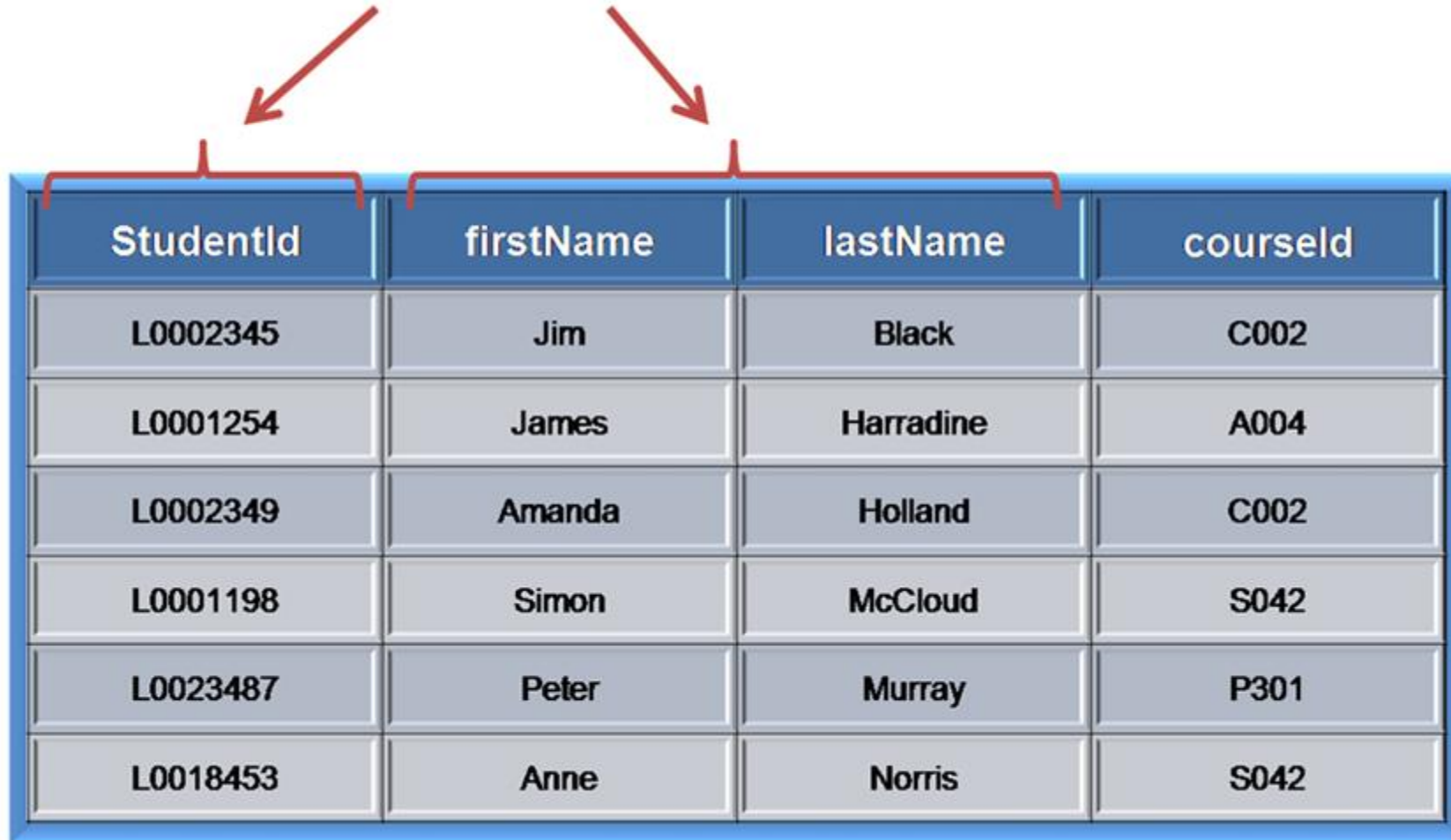


Пример ключей 2

StudentId	firstName	lastName	courseId
L0002345	Jim	Black	C002
L0001254	James	Harradine	A004
L0002349	Amanda	Holland	C002
L0001198	Simon	McCloud	S042
L0023487	Peter	Murray	P301
L0018453	Anne	Norris	S042

Пример ключей 2

Candidate Keys



StudentId	firstName	lastName	courseId
L0002345	Jim	Black	C002
L0001254	James	Harradine	A004
L0002349	Amanda	Holland	C002
L0001198	Simon	McCloud	S042
L0023487	Peter	Murray	P301
L0018453	Anne	Norris	S042

Реляционная модель данных

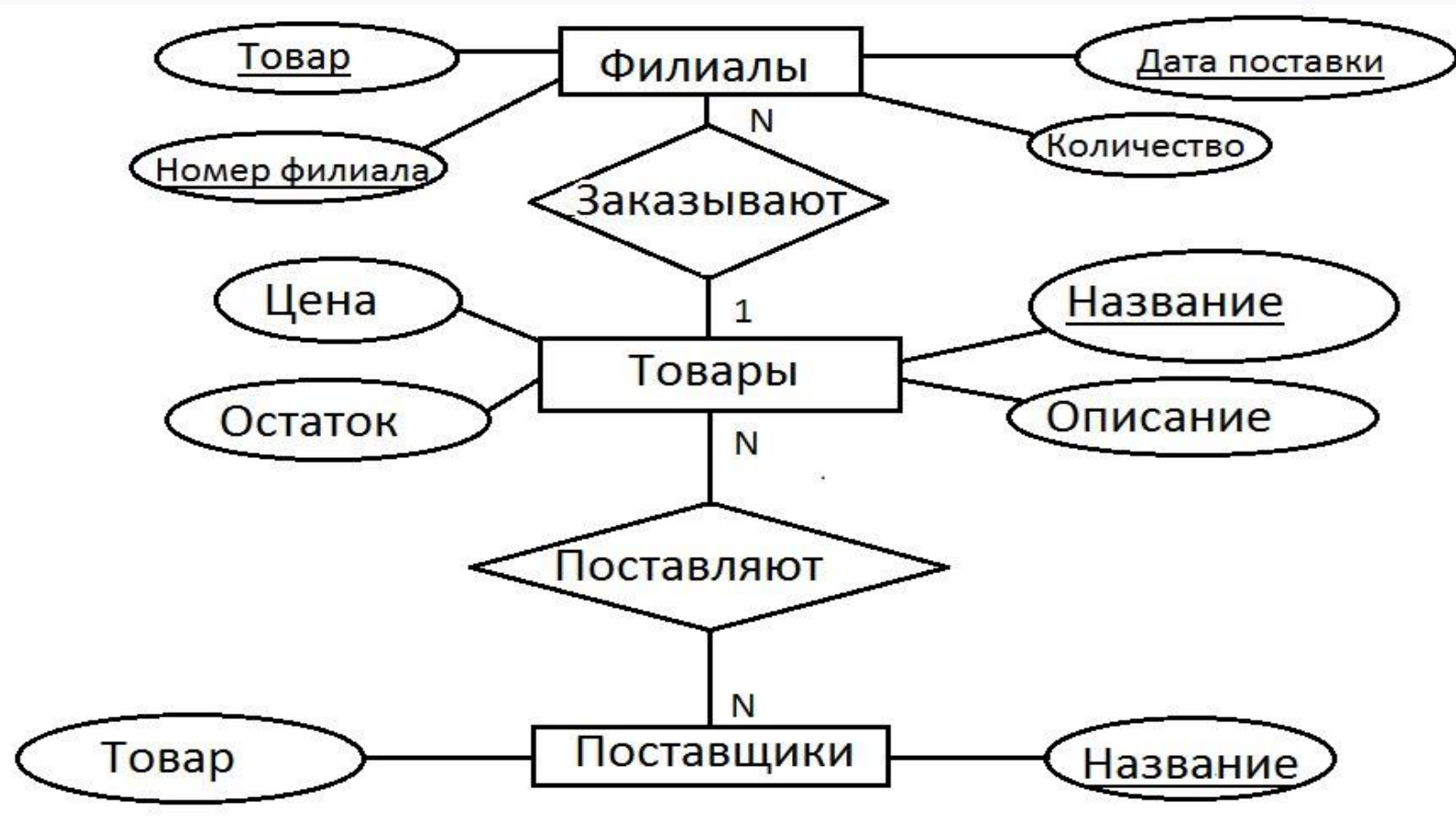
Коллеги, вопросы ?

Этапы проектирования БД

1. Концептуальный уровень.
2. Логический уровень.
3. Физический уровень.

Этапы проектирования БД

1. Концептуальное проектирование - представление данных с точки зрения пользователя.



Результаты концептуального проектирования могут быть представлены в виде модели, которая представляет структуру данных независимо от любой физической реализации.

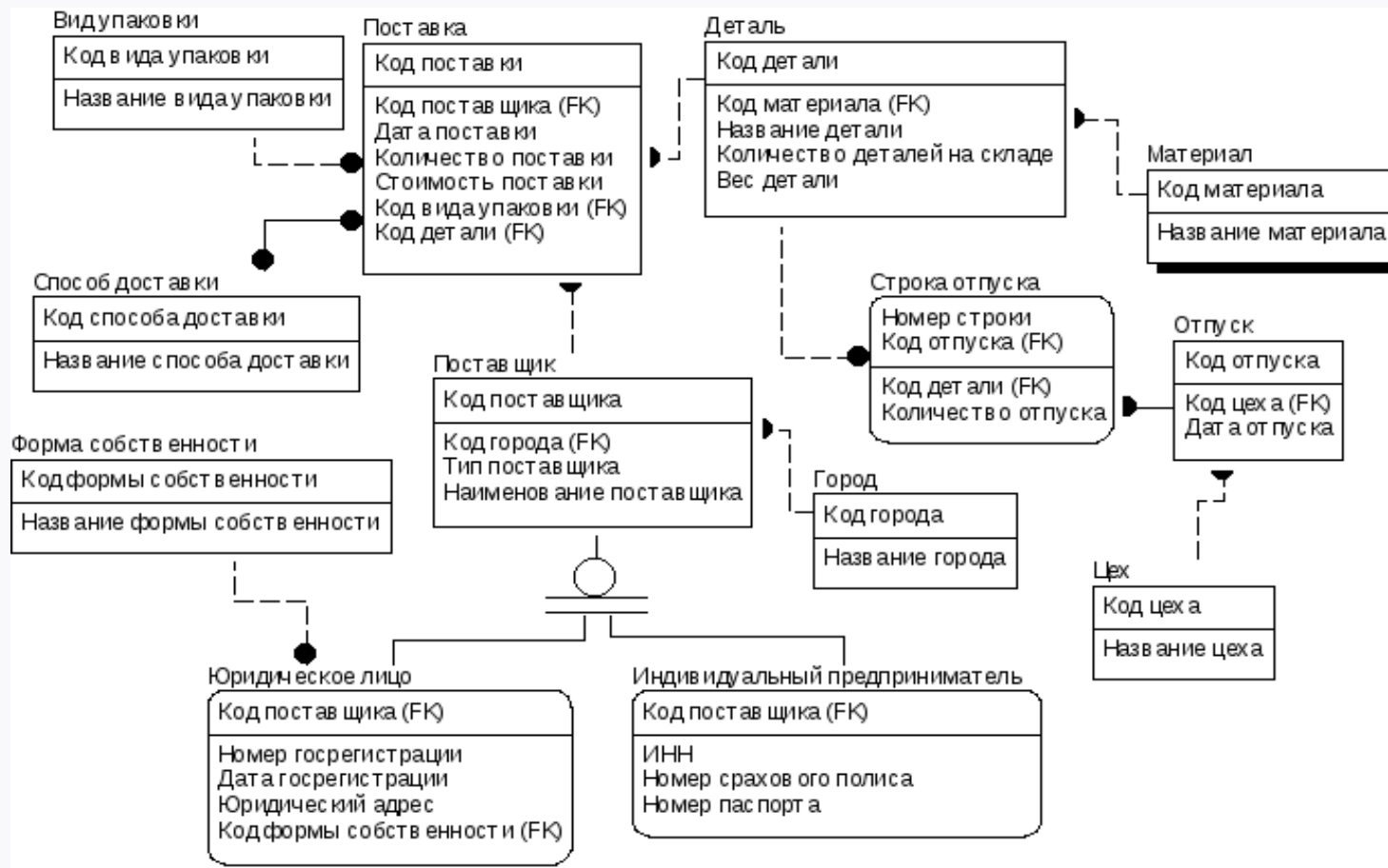
Этапы проектирования БД

Концептуальная модель:

- Определение сущностей и их документирование
- Определение связей между сущностями
- Создание ER модели
- Определение атрибутов
- Определение потенциальных ключей

Этапы проектирования БД

2. Логическое проектирование - представление данных с точки зрения выбранной модели данных.



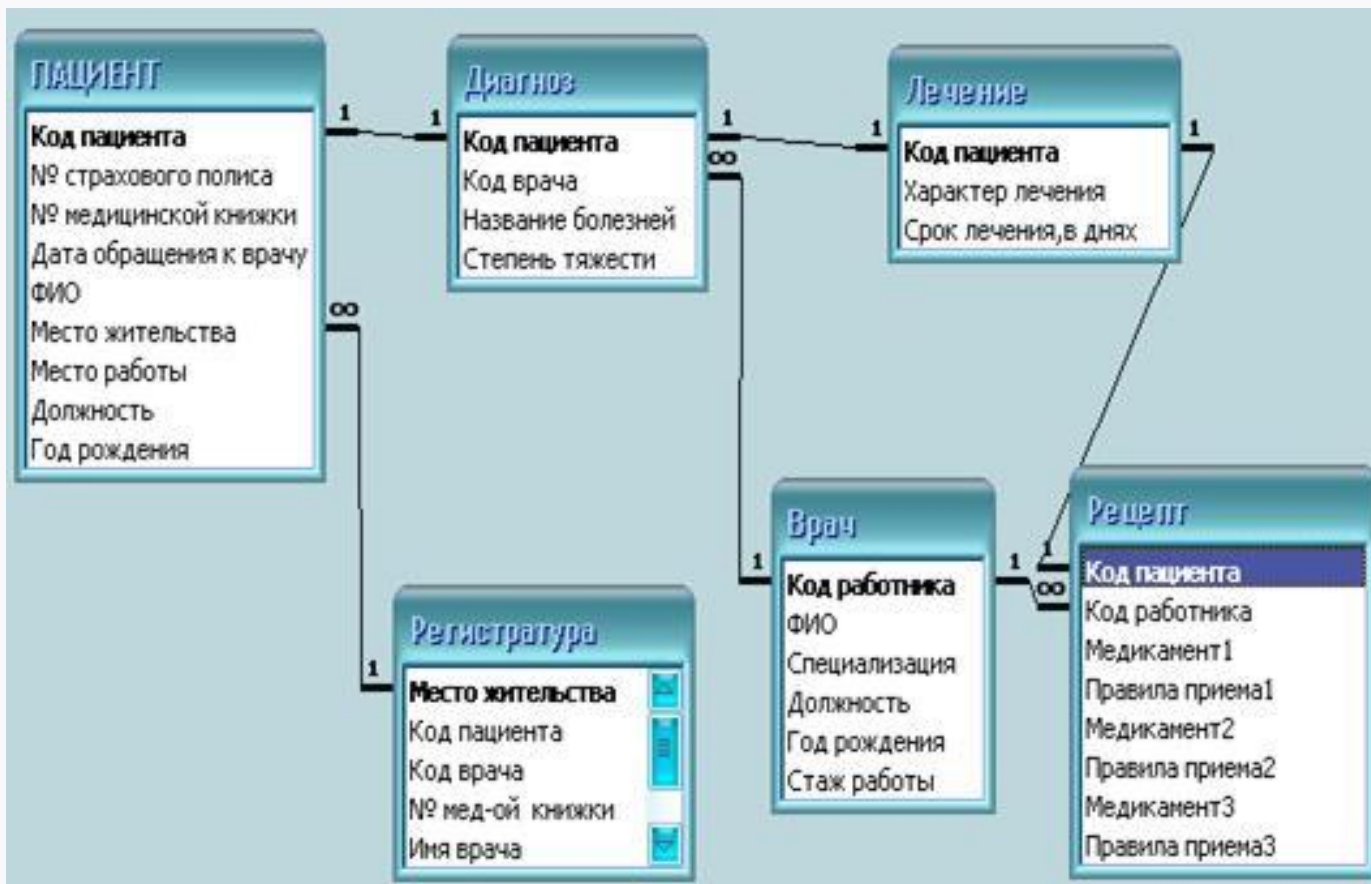
Этапы проектирования БД

Логическая модель:

- Выбор модели данных (РСУБД, постРСУБД)
- Определение набор таблиц
- Нормализация данных
- Денормализация данных
- Определение соответствия структуры данных
- Определение требований поддержки целостности
- Создание окончательной логической модели и обсуждение ее с командой и заказчиками

Этапы проектирования БД

3. Физическое проектирование - представление данных с точки зрения выбранной СУБД.



Этапы проектирования БД

Физическая модель:

- Проектирование таблиц данных средствами выбранной СУБД
- Реализация бизнес правил в выбранной СУБД
- Проектирование физической организации данных
- Планирование ресурсов
- Определение правил безопасности и защиты информации
- Организация сопровождения и мониторинга

Варианты терминов

Концептуальный уровень	Логический уровень	Физический уровень
Сущность	Отношение	Таблица / Файл
Экземпляр	Кортеж	Строка / Запись
Характеристика	Атрибут	Столбец / Поле

Этапы проектирования БД

Коллеги, вопросы ?

Нормализация таблиц

Нормализация - это процесс уменьшения избыточности информации базы данных.

Применяется для устранения:

- **аномалии вставки** - невозможность добавления неполной информации
- **аномалии обновления** - риск неполного обновления
- **аномалии удаления** - риск удаления лишней информации

Нормализация таблиц

Нормализация - это разбиение таблицы на две или больше, обладающих лучшими свойствами при добавлении, изменении и удалении данных.

Цель нормализации - получение такого проекта базы данных, в котором каждый факт появляется лишь в одном месте.

Каждой нормальной форме соответствует некоторый определенный набор ограничений, и отношение находится в некоторой нормальной форме, если удовлетворяет свойственному ей набору ограничений.

Нормализация таблиц

Нормальные формы:

- первая нормальная форма (1NF);
- вторая нормальная форма (2NF);
- третья нормальная форма (3NF);
- нормальная форма Бойса-Кодда (BCNF);
- четвертая нормальная форма (4NF);
- пятая нормальная форма, или нормальная форма проекции-соединения (5NF или PJ/NF).

Основные свойства нормальных форм:

1. Каждая следующая нормальная форма в некотором смысле лучше предыдущей.
2. При переходе к следующей нормальной форме свойства предыдущих нормальных форм сохраняются.

Отношение называется нормализованным или находящимся в 1НФ когда на пересечении строк и столбцов находятся только скалярные (единственные) значения.

1 НФ

CustomerName	Subscription
Пупкин	Ужасы; Документальное кино; Кино о животных
Иванов	Фестивальное кино; Драма
Петров	Комедия; Детектив

CustomerName	Subscription
Пупкин	Ужасы
Пупкин	Документальное кино
Пупкин	Кино о животных
Иванов	Фестивальное кино
Иванов	Драма
Петров	Комедия
Петров	Детектив

1 НФ

CustomerName	Subscription1	Subscription2	Subscription3
Пупкин	Ужасы	Документальное кино	Кино о животных
Иванов	Фестивальное кино	Драма	
Петров	Комедия	Детектив	

CustomerName	Subscription
Пупкин	Ужасы
Пупкин	Документальное кино
Пупкин	Кино о животных
Иванов	Фестивальное кино
Иванов	Драма
Петров	Комедия
Петров	Детектив

1 НФ

CustomerName	CustomerInfo
Пупкин	г. Москва пр. Ленинский 97; pupkin@mail.ru
Иванов	г. Иваново ул. Ленина 5 кв 12; ivanov1990@yandex.ru
Петров	г. Вологда ул Пирожковой 7 кв 112; IvanPetrov@ya.ru

1 НФ

CustomerName	CustomerInfo
Пупкин	г. Москва пр. Ленинский 97; pupkin@mail.ru
Иванов	г. Иваново ул. Ленина 5 кв 12; ivanov1990@yandex.ru
Петров	г. Вологда ул Пирожковой 7 кв 112; IvanPetrov@ya.ru

CustomerName	Address	Email
Пупкин	г. Москва пр. Ленинский 97	pupkin@mail.ru
Иванов	г. Иваново ул. Ленина 5 кв 12	ivanov1990@yandex.ru
Петров	г. Вологда ул Пирожковой 7 кв 112	IvanPetrov@ya.ru

S_id	S_name	town	T_id	T_name	color	qnty
1	Иванов	Москва, МО	50	Ford	Черный	100
			51		Белый	50
2	Петров	Краснодар, КК	52	KIA	Белый	30
			50	Ford	Черный	15
			53	Granta	Серый	100
3	Карпов	Сочи, КК	54	BMW	Синий	20
4	Павлов	Владивосток, ПрК	55	Toyota	Черный	50

S_id	S_name	town	region	T_id	T_name	color	qnty
1	Иванов	Москва	МО	50	Ford	Черный	100
1	Иванов	Москва	МО	51	Ford	Белый	50
2	Петров	Краснодар	КК	52	KIA	Белый	30
2	Петров	Краснодар	КК	50	Ford	Черный	15
2	Петров	Краснодар	КК	53	Granta	Серый	100
3	Карпов	Сочи	КК	54	BMW	Синий	20
4	Павлов	Владивосток	ПрК	55	Toyota	Черный	50
4	Павлов	Владивосток	ПрК	56	Toyota	Белый	25

Функциональные зависимости

Поле В таблицы **функционально зависит** от поля А той же таблицы, когда в любой момент времени для каждого из различных значений поля А обязательно существует только одно из различных значений поля В (поля А и В могут быть составными). $R.A \rightarrow R.B$

Степени функциональной зависимости:

- частичная;
- полная;
- транзитивная;
- многозначная.

Функциональные зависимости

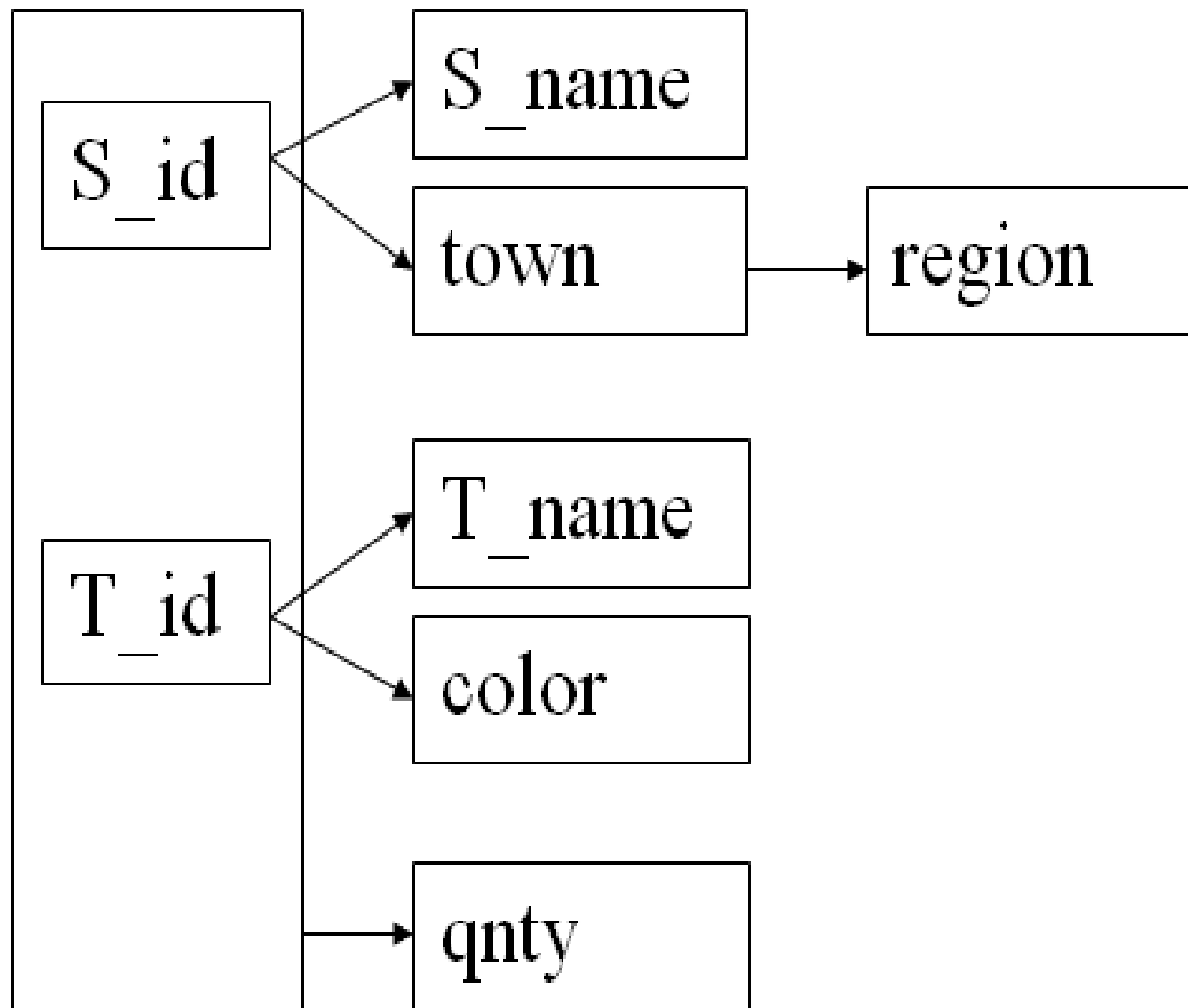
Поле B находится в **полной функциональной зависимости** от составного поля A , если оно функционально зависит от A и не зависит функционально от любого подмножества поля A .

Отношение находится во 2НФ, если оно находится в 1 НФ и каждый неключевой атрибут полностью функционально зависит от первичного ключа.

Что является первичным ключом ???

S_id	S_name	town	region	T_id	T_name	color	qnty
1	Иванов	Москва	МО	50	Ford	Черный	100
1	Иванов	Москва	МО	51	Ford	Белый	50
2	Петров	Краснодар	КК	52	KIA	Белый	30
2	Петров	Краснодар	КК	50	Ford	Черный	15
2	Петров	Краснодар	КК	53	Granta	Серый	100
3	Карпов	Сочи	КК	54	BMW	Синий	20
4	Павлов	Владивосток	ПрК	55	Toyota	Черный	50
4	Павлов	Владивосток	ПрК	56	Toyota	Белый	25

S_id	S_name	town	region	T_id	T_name	color	qnty
1	Иванов	Москва	МО	50	Ford	Черный	100
1	Иванов	Москва	МО	51	Ford	Белый	50
2	Петров	Краснодар	КК	52	KIA	Белый	30
2	Петров	Краснодар	КК	50	Ford	Черный	15
2	Петров	Краснодар	КК	53	Granta	Серый	100
3	Карпов	Сочи	КК	54	BMW	Синий	20
4	Павлов	Владивосток	ПрК	55	Toyota	Черный	50
4	Павлов	Владивосток	ПрК	56	Toyota	Белый	25



S_id	S_name	town	region
1	Иванов	Москва	МО
1	Иванов	Москва	МО
2	Петров	Краснодар	КК
2	Петров	Краснодар	КК
2	Петров	Краснодар	КК
3	Карпов	Сочи	КК
4	Павлов	Владивосто к	ПрК
4	Павлов	Владивосто к	ПрК

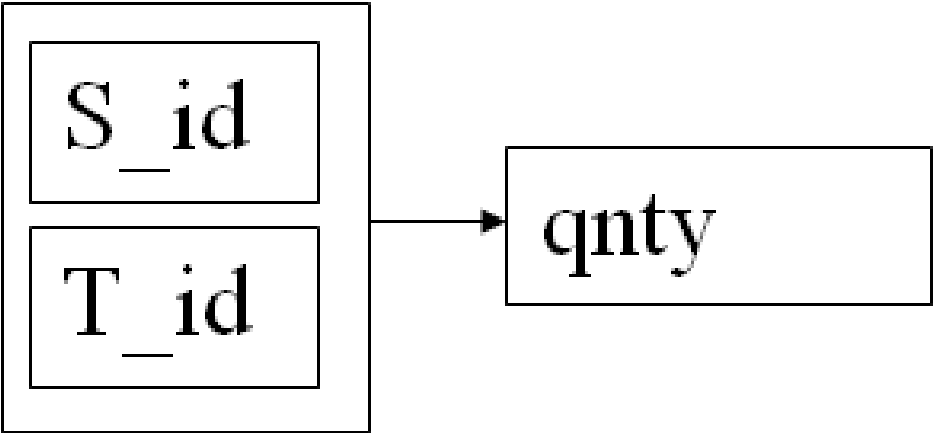
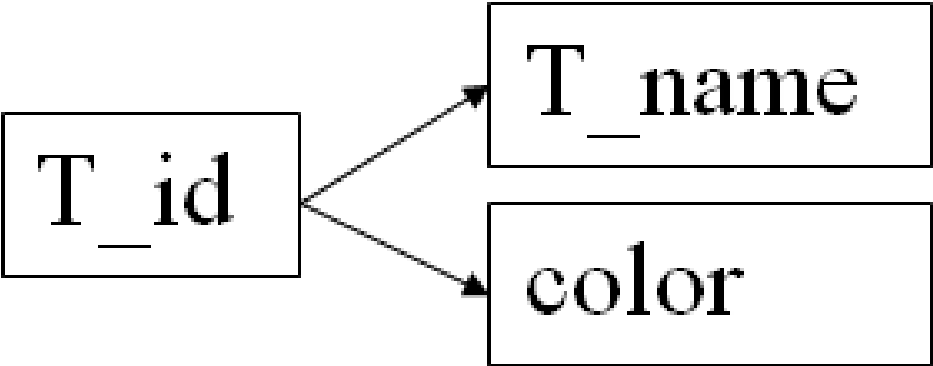
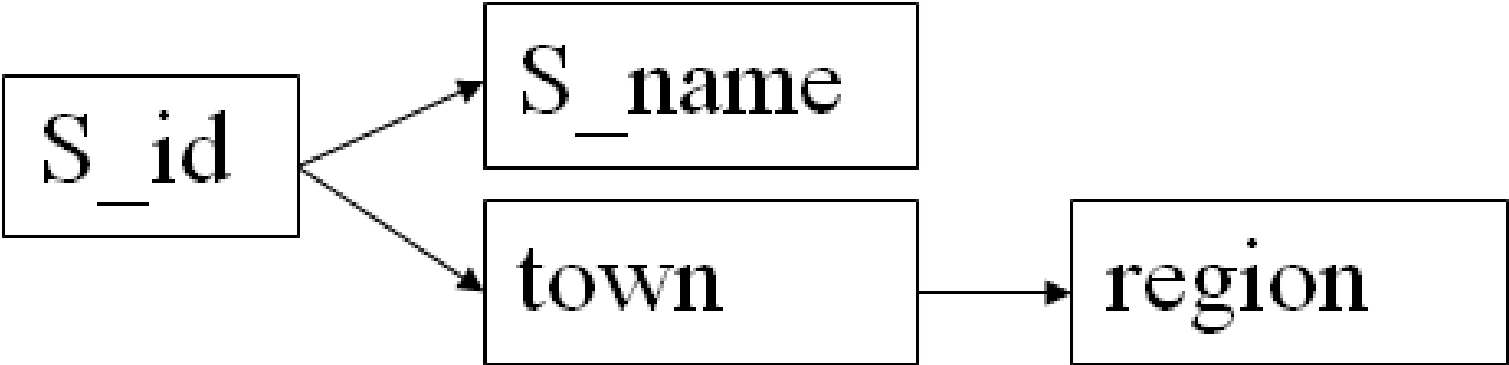
T_id	T_name	color
50	Ford	Черный
51	Ford	Белый
52	KIA	Белый
50	Ford	Черный
53	Granta	Серый
54	BMW	Синий
55	Toyota	Черный
56	Toyota	Белый

S_id	T_id	qnty
1	50	100
1	51	50
2	52	30
2	50	15
2	53	100
3	54	20
4	55	50
4	56	25

S_id	S_name	town	region
1	Иванов	Москва	МО
2	Петров	Краснодар	КК
3	Карпов	Сочи	КК
4	Павлов	Владивосто к	ПрК

T_id	T_name	color
50	Ford	Черный
51	Ford	Белый
52	KIA	Белый
53	Granta	Серый
54	BMW	Синий
55	Toyota	Черный
56	Toyota	Белый

S_id	T_id	qnty
1	50	100
1	51	50
2	52	30
2	50	15
2	53	100
3	54	20
4	55	50
4	56	25



Функциональные зависимости

Функциональная зависимость $R.A \rightarrow R.B$ называется **транзитивной**, если существует такой атрибут C , что имеются функциональные зависимости $R.A \rightarrow R.C$ и $R.C \rightarrow R.B$ и отсутствует функциональная зависимость $R.C \rightarrow R.A$.

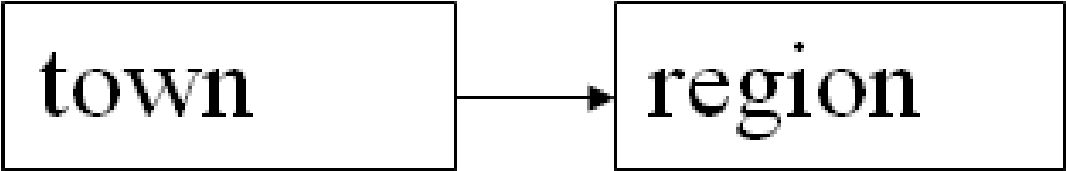
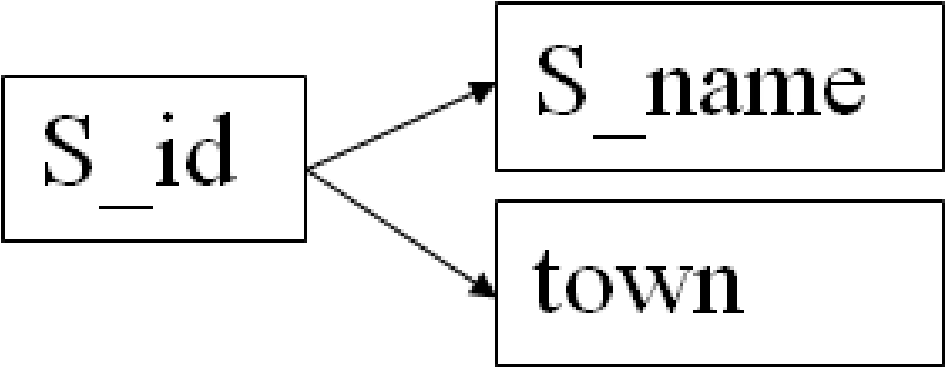
3 НФ

Отношение находится в 3НФ, когда находится во 2НФ и каждый неключевой атрибут нетранзитивно зависит от первичного ключа.

Отношение находится в 3НФ, если оно находится во 2НФ и каждый неключевой атрибут не зависит от другого неключевого атрибута.

S_id	S_name	town
1	Иванов	Москва
2	Петров	Краснодар
3	Карпов	Сочи
4	Павлов	Владивосток

town	region
Москва	МО
Краснодар	КК
Сочи	КК
Владивосток	ПрК



S_id	S_name	town
1	Иванов	Москва
2	Петров	Краснодар
3	Карпов	Сочи
4	Павлов	Владивосток

town	region
Москва	МО
Краснодар	КК
Сочи	КК
Владивосток	ПрК

S_id	T_id	qnty
1	50	100
1	51	50
2	52	30
2	50	15
2	53	100
3	54	20
4	55	50
4	56	25

T_id	T_name	color
50	Ford	Черный
51	Ford	Белый
52	KIA	Белый
53	Granta	Серый
54	BMW	Синий
55	Toyota	Черный
56	Toyota	Белый

НФ Бойса-Кодда (BCNF)

Детерминант - любой атрибут, от которого полностью функционально зависит некоторый другой атрибут (левая часть функциональной зависимости $A \rightarrow B$).

Таблица приведена к третьей нормальной форме Бойса--Кодда, когда детерминанты всех ее функциональных зависимостей являются потенциальными ключами.

Применяется:

1. Отношение имеет две или более потенциальных ключа.
2. Два и более потенциальных ключа являются составными.
3. Ключи пересекаются, т.е. имеют хотя бы один общий атрибут.

<u>CustomerName</u>	Subscription	Author
Пупкин	Кино о животных	Дроздов
Иванов	Фестивальное кино	Хрустов
Иванов	Драма	Сидоров
Петров	Комедия	Хрустов
Петров	Детектив	Сидоров

<u>CustomerName</u>	Subscription	Author
Пупкин	Кино о животных	Дроздов
Иванов	Фестивальное кино	Хрустов
Иванов	Драма	Сидоров
Петров	Комедия	Хрустов
Петров	Детектив	Сидоров

CustomerName	Subscription
Пупкин	Кино о животных
Иванов	Фестивальное кино
Иванов	Драма
Петров	Комедия
Петров	Детектив

Author	Subscription
Дроздов	Кино о животных
Хрустов	Фестивальное кино
Сидоров	Драма
Хрустов	Комедия
Сидоров	Детектив

Subscription
Кино о животных
Фестивальное кино
Драма
Комедия
Детектив

Отношение находится в 4 НФ, если она соответствует 3NF и в ней отсутствуют многозначные зависимости.

Многозначная зависимость имеет место в ситуации, когда одному значению атрибута А соответствует несколько значений атрибута В.

S_name	T_name	Hobby
Иванов	Ford	Футбол
Иванов	Ford	Бег
Петров	KIA	Футбол
Петров	Ford	Футбол
Карпов	BMW	NULL
Павлов	NULL	Футбол
Сидоров	NULL	NULL

PR: {S_name, T_name, Hobby}

Многозначная зависимость:
S_name ->-> T_name, Hobby

S_name
Иванов
Петров
Карпов
Павлов
Сидоров

S_name	T_name
Иванов	Ford
Петров	KIA
Петров	Ford
Карпов	BMW
Павлов	NULL
Сидоров	NULL

S_name	Hobby
Иванов	Футбол
Иванов	Бег
Петров	Футбол
Карпов	NULL
Павлов	Футбол
Сидоров	NULL

Отношение находится в 5 НФ тогда и только тогда, когда каждая нетривиальная зависимость соединения подразумевается ее ключами.

Поставщик	Товар	Потребитель
Ф 1	Товар 1	П 2
Ф 1	Товар 2	П 1
Ф 2	Товар 1	П 1
Ф 1	Товар 1	П 1

ПТ

Постав- щик	Товар
Ф 1	Товар 1
Ф 1	Товар 2
Ф 2	Товар 1

ТП

Товар	Потре- битель
Товар 1	П 2
Товар 2	П 1
Товар 1	П 1

ПШ

Постав- щик	Потре- битель
Ф 1	П 2
Ф 1	П 1
Ф 2	П 1

Нормализация таблиц

Коллеги, вопросы ?

Рефлексия

О чем мы сегодня говорили?

- Сколько этапов проектирования БД?
- Какие аномалии вы запомнили?
- Что такое нормализация и до какой формы обычно рекомендуется проводить нормализацию?

The background of the entire image is an aerial photograph of a city with many skyscrapers, overlaid with a semi-transparent blue layer. A network of thin, light-blue lines connects various points across the blue area, creating a digital or technological aesthetic.

Заполните, пожалуйста,
опрос о занятии по ссылке в чате

Спасибо за внимание!
Приходите на следующие вебинары



Коробков Виктор