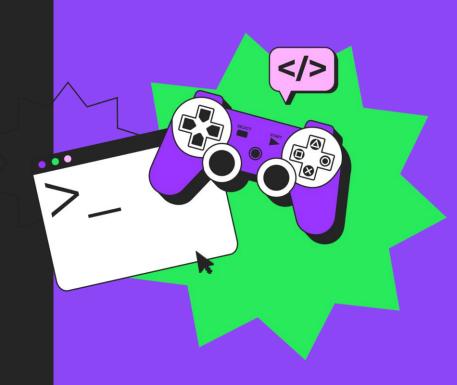


Получение денормализованных таблиц из нормализованных.

Урок 3

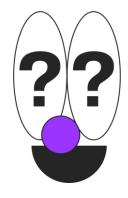




### Ответьте на несколько вопросов сообщением в чат



Из какого вы города?



Сколько вам лет?



Кем вы работаете сейчас? Как долго?



#### План курса (вертикальный)

Обзор возможностей Airflow, установка и настройка. Модели данных и нормализация таблиц. Схема "звезда". Операторы в Airflow и их применение для ETL. Введение в подготовку данных для аналитиков. Таблицы фактов и таблицы измерений. Построение пайплайнов и визуализация потоков данных в Airflow. Получение денормализованных таблиц из нормализованных. Специфика применения ETL в различных предметных сферах Партицирование данных по дате. Динамическое партицирование.



#### Что будет на уроке сегодня

- 於 Что такое денормализация и зачем
- 於 Когда полезно использовать денормализацию
- 於 Как определить что денормализация оправдана
- 於 Как грамотно реализовать денормализацию



### Викторина



#### Для чего чаще всего применяется денормализация?

- 1. Ускорение операций чтения
- 2. Уменьшение количества логических ошибок
- 3. Минимизирует избыточность данных



#### Для чего чаще всего применяется денормализация?

- 1. Ускорение операций чтения
- 2. Уменьшение количества логических ошибок
- 3. Минимизирует избыточность данных



#### Когда стоит задуматься о денормализации?

- 1. Необходимо сохранять исторические данные
- 2. Повышение производительности запросов
- 3. Предварительные вычисления часто запрашиваемых значений
- 4. Все варианты верны



#### Когда стоит задуматься о денормализации?

- 1. Необходимо сохранять исторические данные
- 2. Повышение производительности запросов
- 3. Предварительные вычисления часто запрашиваемых значений
- 4. Все варианты верны



## Необходима ли денормализация если мы часто обращаемся к таблице содержащей длинное поле?

- 1. Обязательно, это сильно ускорит выполнение запросов
- 2. Необязательно, сильного прироста производительности ожидать не следует
- 3. Нет не нужно



## Необходима ли денормализация если мы часто обращаемся к таблице содержащей длинное поле?

- 1. Обязательно, это сильно ускорит выполнение запросов
- 2. Необязательно, сильного прироста производительности ожидать не следует
- 3. Нет не нужно



#### Два основных метода денормализации данных?

- 1. Дублирование и репликация
- 2. Предварительная подготовка и очистка
- 3. Дублирование и предварительная подготовка
- 4. Предварительная подготовка и репликация



#### Два основных метода денормализации данных?

- 1. Дублирование и репликация
- 2. Предварительная подготовка и очистка
- 3. Дублирование и предварительная подготовка
- 4. Предварительная подготовка и репликация









# Вопросы?











### Практика

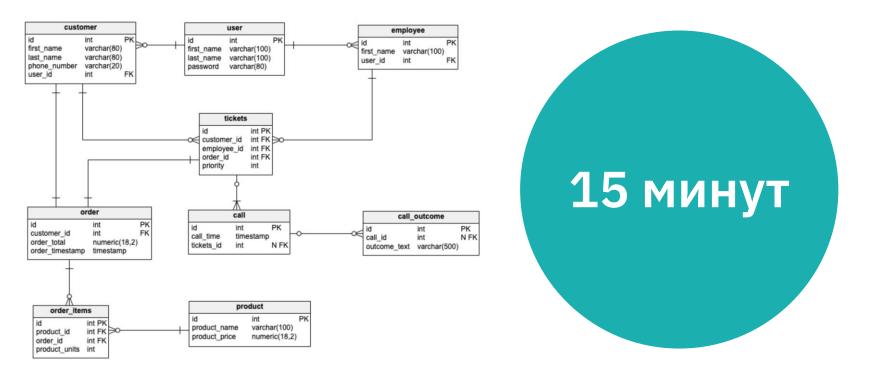


#### Денормализация данных

**Денормализация** (англ. denormalization) — намеренное приведение структуры базы данных в состояние, не соответствующее критериям нормализации, обычно проводимое с целью ускорения операций чтения из базы за счет добавления избыточных данных.



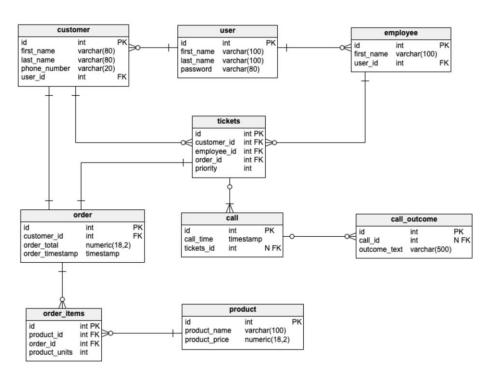
Ниже представлена схема базы данных приложения в котором клиент может размещать заказы и открывать тикеты для службы поддержки по своим заказам. Посмотрите схему базы данных, опишите структуру и дату которая хранится в этих таблицах.

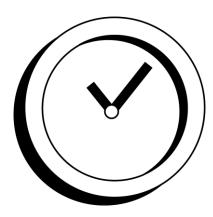






Посмотрите схему базы данных, опишите структуру и дату которая хранится в этих таблицах







Предположим что у нас появилось две новых бизнес-задачи.

Во-первых, каждому представителю службы поддержки нужен быстрый доступ к некоторым показателям клиента. Для новой функциональности в приложении требуется информационная панель о клиенте с общим объемом продаж, количеством отправленных билетов и количеством обращений в службу поддержки клиентов.

Во-вторых, представителям службы поддержки нужна панель управления тиектами со списком тикетов и их соответствующими данными. Это нужно им для выбора тикета по адресу в зависимости от приоритета.

Создайте таблицу customer\_statistics которая будет содержать информацию о покупательских привычках каждого клиента.

Сделайте схему таблицы используя <u>app.diagrams.net</u> и поделитесь картинкой в чате





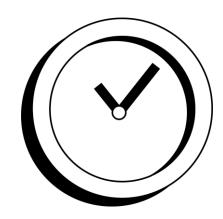
<<15:00->> Предположим что у нас появилось две новых бизнес-задачи.

Во-первых, каждому представителю службы поддержки нужен быстрый доступ к некоторым показателям клиента. Для новой функциональности в приложении требуется информационная панель о клиенте с общим объемом продаж, количеством отправленных билетов и количеством обращений в службу поддержки клиентов.

Во-вторых, представителям службы поддержки нужна панель управления тикетами со списком тикетов и их соответствующими данными. Это нужно им для выбора тикета по адресу в зависимости от приоритета.

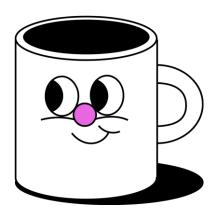
Создайте таблицу customer statistics которая будет содержать информацию о покупательских привычках каждого клиента.

Сделайте схему таблицы используя app.diagrams.net и поделитесь картинкой в чате





#### Перерыв



<<5:00->>



Измените таблицу tickets таким образом чтобы она соответствовала требованиям из предыдущего задания.

Представителям службы поддержки нужна панель управления тикетами со списком тикетов и их соответствующими данными. Это нужно им для выбора тикетоа по адресу в зависимости от приоритета.

Сделайте схему таблицы используя <u>app.diagrams.net</u> и поделитесь картинкой в чате



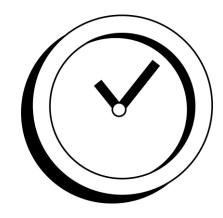


<<15:00->>

Измените таблицу tickets таким образом чтобы она соответствовала требованиям из предыдущего задания.

Представителям службы поддержки нужна панель управления тикетами со списком тикетов и их соответствующими данными. Это нужно им для выбора тикетов по адресу в зависимости от приоритета.

Сделайте схему таблицы используя <u>app.diagrams.net</u> и поделитесь картинкой в чате





Представитель службы поддержки заметил, что сумма, уплаченная за продукт, не соответствует цене, указанной в дашборде. Это произошло потому что с момента оплаты клиентом товара цена изменилась.

Подумайте что нужно изменить в исходной схеме (не создавая новых таблиц) чтобы решить описанную выше проблему.

Hapucyйте схему измененных таблиц используя app.diagrams.net и поделитесь картинкой в чате.



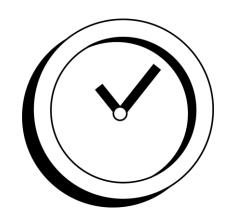


<<15:00->>

Представитель службы поддержки заметил, что сумма, уплаченная за продукт, не соответствует цене, указанной в дашборде. Это произошло потому что с момента оплаты клиентом товара цена изменилась.

Подумайте что нужно изменить в исходной схеме (не создавая новых таблиц) чтобы решить описанную выше проблему.

Hapucyйте схему измененных таблиц используя <u>app.diagrams.net</u> и поделитесь картинкой в чате.











# Вопросы?









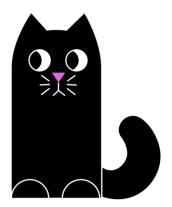


### Домашнее задание



#### Домашнее задание

текст задания





### Спасибо за внимание

