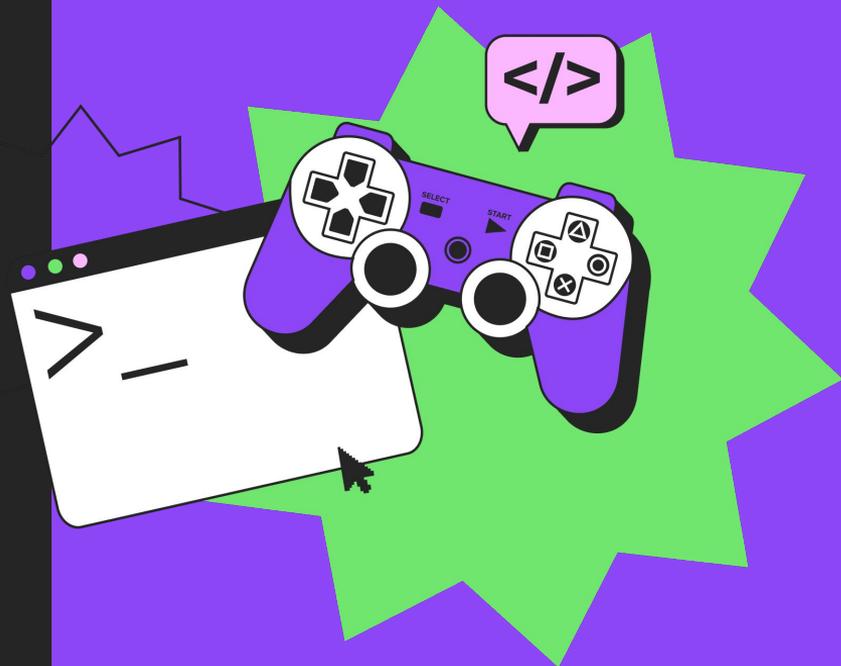


# Получение денормализованных таблиц из нормализованных .

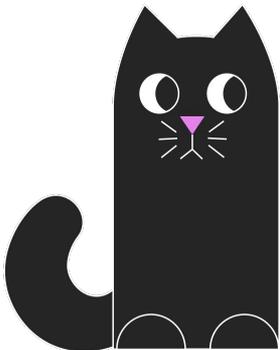
Урок 3



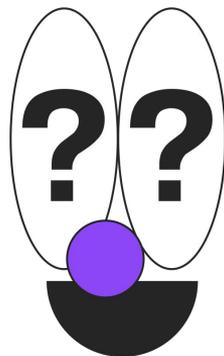
Теперь ваша очередь!



## Ответьте на несколько вопросов сообщением в чат



**Из какого вы города?**



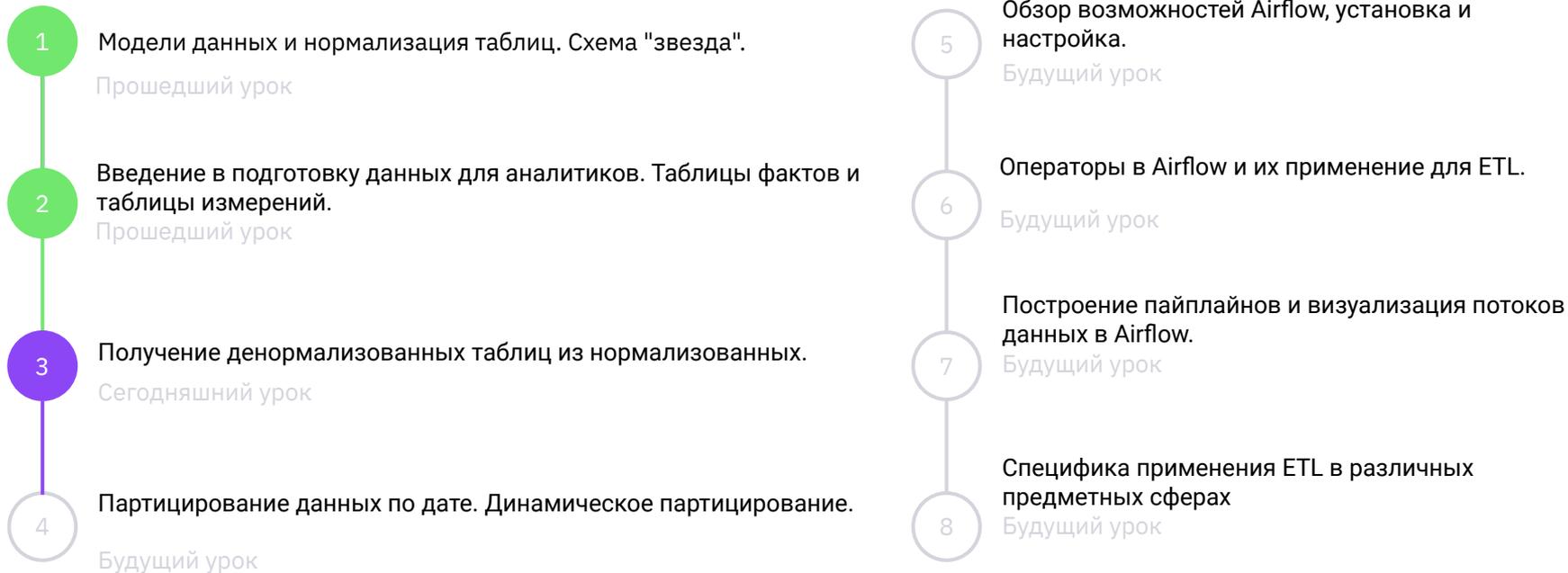
**Сколько вам лет?**



**Кем вы работаете сейчас?  
Как долго?**



## План курса (вертикальный)





## Что будет на уроке сегодня

-  Что такое денормализация и зачем
-  Когда полезно использовать денормализацию
-  Как определить что денормализация оправдана
-  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?



Вопросы?



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.

# Задание 0



Заказчик попросил разработать таблицу для BI отчета, благодаря которой можно посчитать длительность статуса и время на группы и понять к какой группе относится тот или иной статус, построить статусную-временную модель в PBI.

Нужно:

Привести таблицу протокола из этого вида в этот вид.

objectid	restime	desc	fieldvalue	fieldname
15	1 653 328 726	'Статус' изменен на:...	В работе	Status
15	1 653 364 688	'Группа исполнителя'...	ЭЛС\Север	GNAME2
15	1 653 364 688	'Статус' изменен на:...	Назначен	Status
15	1 653 365 165	'Статус' изменен на:...	В работе	Status
15	1 653 365 268	'Статус' изменен на:...	Решен	Status
15	1 653 365 279	'Статус' изменен на:...	Исследование ситуа...	Status
15	1 653 365 287	'Статус' изменен на:...	Закрыт	Status
16	1 659 305 822	Зона ответственности...	МГТС	Resp_zone
16	1 659 305 822	'Статус' изменен на:...	Зарегистрирован	Status
16	1 659 305 922	Группа исполнителя'...	ОТЭ	GNAME2
16	1 659 305 932	'Статус' изменен на:...	Назначен	Status
16	1 659 306 922	'Статус' изменен на:...	В работе	Status
16	1 659 306 922	Группа исполнителя'...	ОТПМР	GNAME2
16	1 659 306 922	Зона ответственности...	Клиент	Resp_zone
16	1 659 310 922	Группа исполнителя'...	ЭЛС	GNAME2
16	1 659 311 922	'Статус' изменен на:...	Решен	Status
16	1 659 311 950	Группа исполнителя'...	ОТЭ	GNAME2
16	1 659 311 950	'Статус' изменен на:...	Закрыт	Status

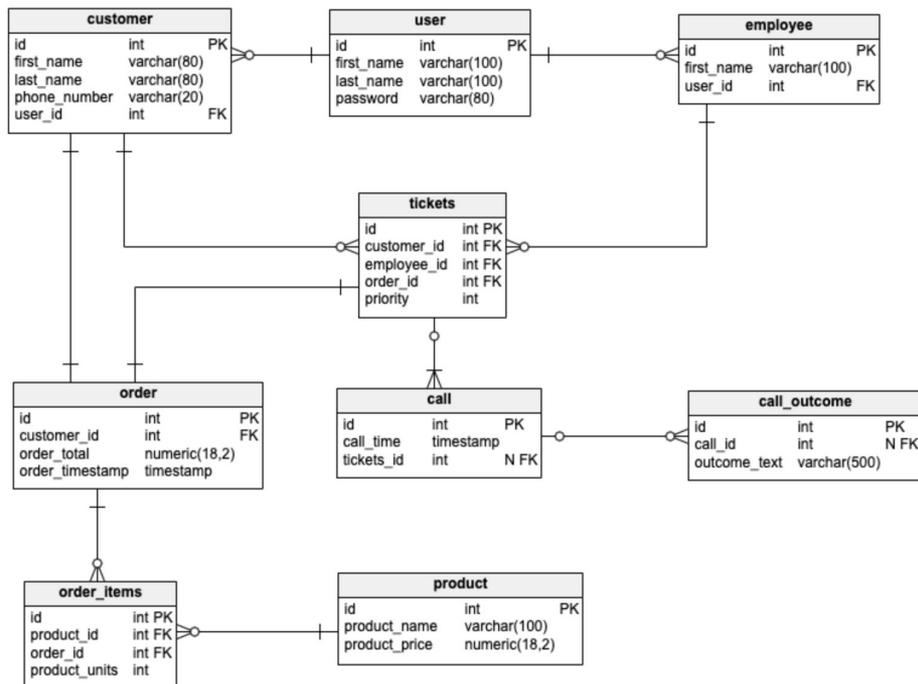
ID_тикета	StatusTime	Длитель...	Статус	Группа	Назначения
15	2022-05-23 14:37:02	0,0003	Зарегистрирован	ЦМик\СДИА	1
15	2022-05-23 14:37:03	0,0861	Назначен	ЦМик\СДИА	(NULL)
15	2022-05-23 14:42:13	3,9072	В работе	ЦМик\СДИА	(NULL)
15	2022-05-23 18:36:39	0,6011	Назначен	ЭЛС\АВР	1
15	2022-05-23 19:12:43	0,1508	В работе	ЭЛС\АВР	(NULL)
15	2022-05-23 19:21:46	0,325	Назначен	ЦАВР\Р\Ф\Г\К\АВР	1
15	2022-05-23 19:41:16	0,0033	Назначен	ЦАВР\Р\ОСПД\ГР...	1
15	2022-05-23 19:41:28	1,2286	В работе	ЦАВР\Р\ОСПД\ГР...	(NULL)
15	2022-05-23 20:55:11	0,0222	Назначен	ЦМик\СДИА	1
15	2022-05-23 20:56:31	0,0375	Назначен	ЭЛС\АВР	1
15	2022-05-23 20:56:46	9,9894	В работе	ЭЛС\АВР	(NULL)
15	2022-05-24 06:58:08	0,1325	Назначен	ЭЛС\Север	1
15	2022-05-24 07:06:05	0,0286	В работе	ЭЛС\Север	(NULL)
15	2022-05-24 07:07:48	0,0031	Решен	ЭЛС\Север	(NULL)
15	2022-05-24 07:07:59	0,0022	Исследование ситуац...	ЭЛС\Север	(NULL)
15	2022-05-24 07:08:07	(NULL)	Закрыт	ЭЛС\Север	(NULL)
16	2022-08-01 01:17:02	0,0278	Зарегистрирован	(NULL)	(NULL)
16	2022-08-01 01:18:42	0,0028	Зарегистрирован	ОТЭ	1
16	2022-08-01 01:18:52	0,275	Назначен	ОТЭ	(NULL)
16	2022-08-01 01:35:22	1,1111	В работе	ОТПМР	1
16	2022-08-01 02:42:02	0,2778	В работе	ЭЛС	1
16	2022-08-01 02:58:42	0,0078	Решен	ЭЛС	(NULL)
16	2022-08-01 02:59:10	(NULL)	Закрыт	ОТЭ	1

Для этого напиши ETL на spark, который перенесет данные из s3.xlsx в mysql и создаст запросом еще одну таблицу в требуемом виде.



## Задание 1

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.



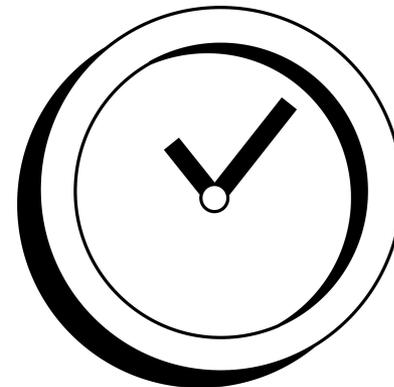
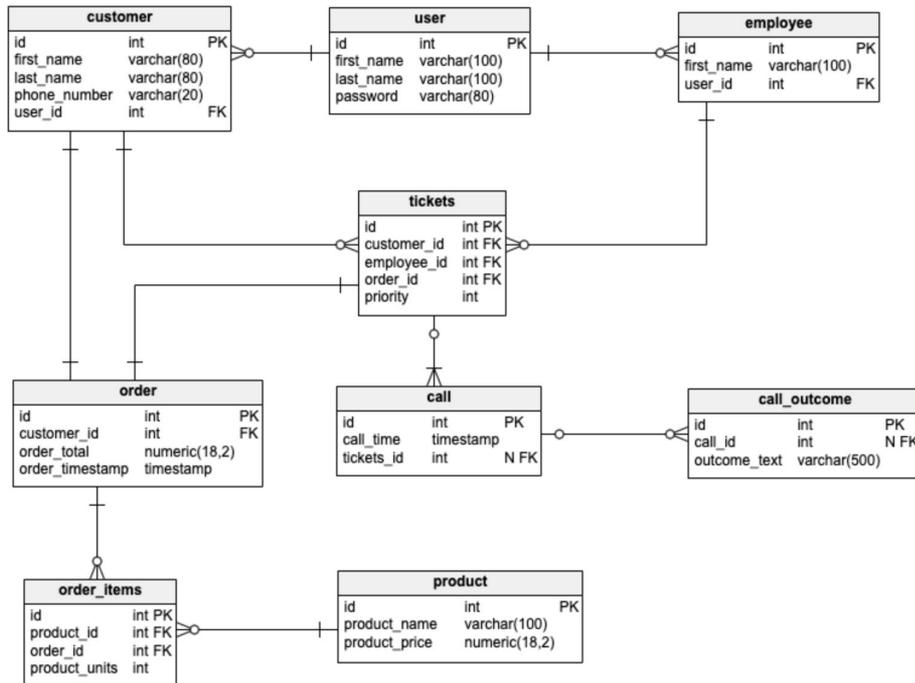
15 минут



## Задание 1

# <<15:00->>

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





## Задание 2

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](https://app.diagrams.net) и поделитесь картинкой в чате



15 минут



## Задание 2

<<15:00->>

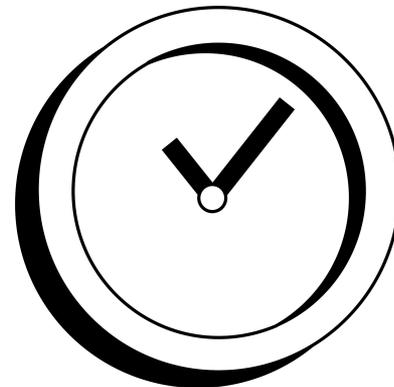
Предположим что у нас появилось две новых бизнес-задачи.

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.

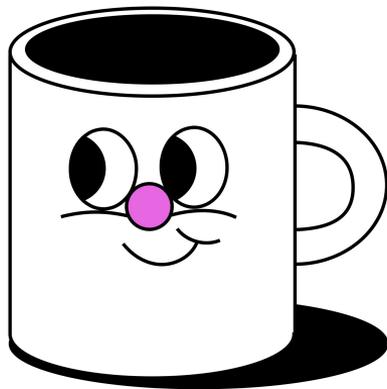
Создайте таблицу `customer_statistics` которая будет содержать информацию о покупательских привычках каждого клиента.

Сделайте схему таблицы используя [app.diagrams.net](https://app.diagrams.net) и поделитесь картинкой в чате





## Перерыв



<<5:00->>



## Задание 3

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 тикета по адресу в зависимости от приоритета.

Сделайте схему таблицы используя [app.diagrams.net](https://app.diagrams.net) и поделитесь картинкой в чате



15 минут



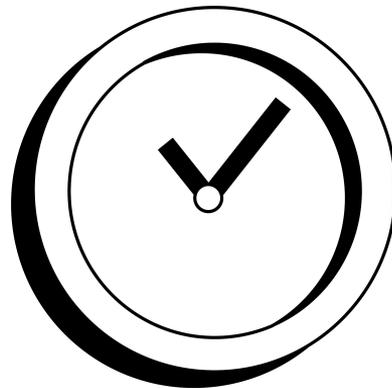
## Задание 3

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 [app.diagrams.net](https://app.diagrams.net) и поделитесь картинкой в чате

<<15:00->>





## Задание 4

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.

Нарисуйте схему измененных таблиц используя [app.diagrams.net](https://app.diagrams.net) и поделитесь картинкой в чате.



**15 минут**



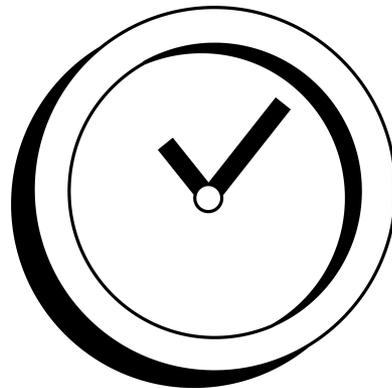
## Задание 4

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.

Нарисуйте схему измененных таблиц используя [app.diagrams.net](https://app.diagrams.net) и поделитесь картинкой в чате.

<<15:00->>





Вопросы?



Вопросы?



Вопросы?





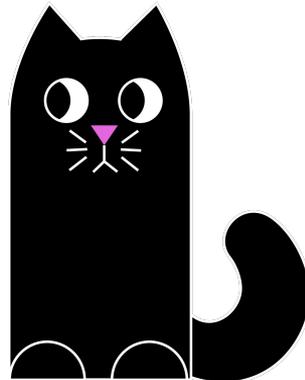
# Домашнее задание



## Домашнее задание

На основе сайта [yandex.ru](https://yandex.ru):

- Определите, на каком протоколе работает сайт.
- Проанализируйте структуру страницы сайта
- Внесите не менее 10 изменений на страницу с помощью инструмента разработчика и представьте скриншоты было/стало.
- Создайте прототип низкой детализации (дополнительное задание, если на семинаре дошли до задания №8)





Спасибо за внимание

