**Изучите данные и выполните следующие задачи. Выполните задание с помощью SQL.**

**Решение оформите в одном Word-файле с текстом запросов на каждое задание.**

Есть два набора данных: визиты пользователей на сервис и данные по покупкам. У каждого визита есть источник, а также устройство, с которого заходил пользователь.

Скачайте файл с данными и выполните задачи на соответствующих страницах

Таблица ***visits***

- *uid* (идентификатор пользователя);

- *source\_id* (id источника визита);

- *device* (устройство);

- *start\_dttm* (дата и время старта визита);

- *end\_dttm* (дата и время окончания визита).

Таблица ***purchases***

- *uid* (идентификатор пользователя);

- *buy\_ts* (дата и время покупки);

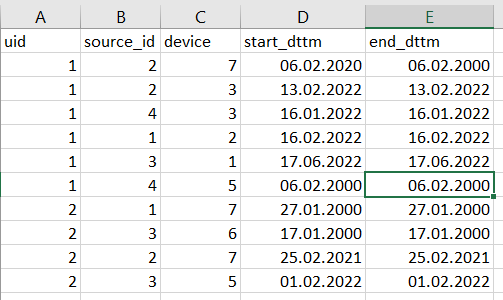
- *revenue* (доход в y. e., клиентский чек).

*Обратите внимание, что в Metabase данных таблиц нет. Выдача заданий с «вымышленными» таблицами — это частая практика на собеседованиях и ТЗ от работодателей.*

*Чтобы протестировать свой код на корректную работу, можно создать таблицы с данными на ресурсе* <https://sqliteonline.com/> *и отработать запросы.*

**Напишите запросы для следующих задач:**

В экселе созданы таблицы visits и ***purchases для тестирования на портале*** <https://sqliteonline.com/>



***Задание 2.1***

*Постройте количество визитов в динамике по месяцам старта визита. Ограничьтесь только визитами 2022 года.*

-- выбрали месяцы и счет построчно

Мне пришлось дополнительно изменять в MySql формат дат и работать уже с временной таблицей

*WITH date1 as*

*(SELECT \*,*

*to\_date(start\_dttm, 'dd/MM/yyyy') as month1*

*from visits)*

*SELECT*

*DATE\_TRUNC('month', month1) AS month,*

*COUNT(\*) AS visit\_count*

*FROM*

*date1*

*WHERE*

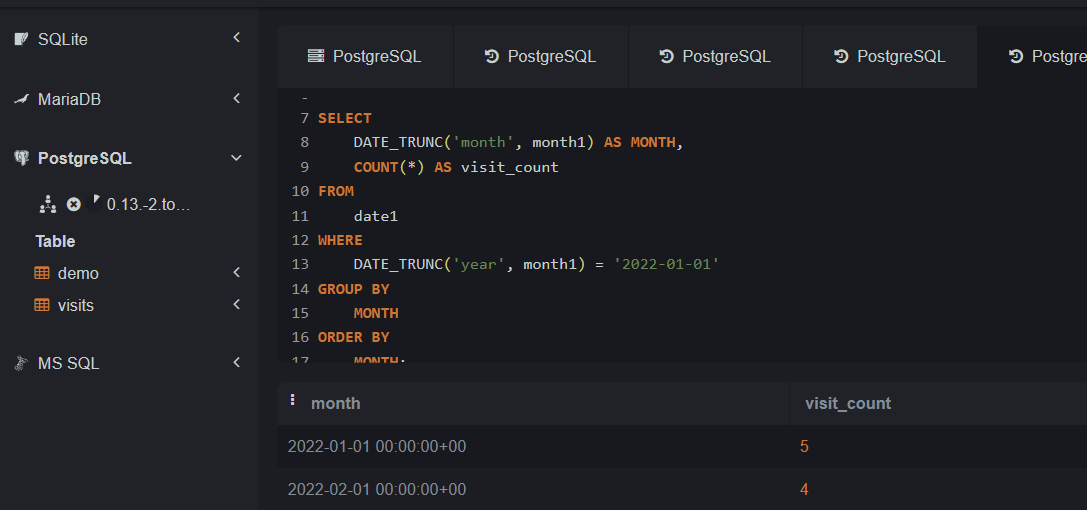
*DATE\_TRUNC('year', month1) = '2022-01-01'*

*GROUP BY*

*month*

*ORDER BY*

*month;*

**

***Задание 2.2***

***Выведите устройства, по которым было больше 100 визитов за июнь 2022 года.***

*-- подгоняем дату*

*WITH date1 as*

*(SELECT \*,*

*to\_date(start\_dttm, 'dd/MM/yyyy') as month1*

*from visits)*

*-- Выбираем девайс*

*SELECT device*

*-- из таблицы визитов*

*FROM data1*

*-- за июнь 2022*

*WHERE start\_dttm >= '2022-06-01 00:00:00' AND start\_dttm < '2022-07-01 00:00:00'*

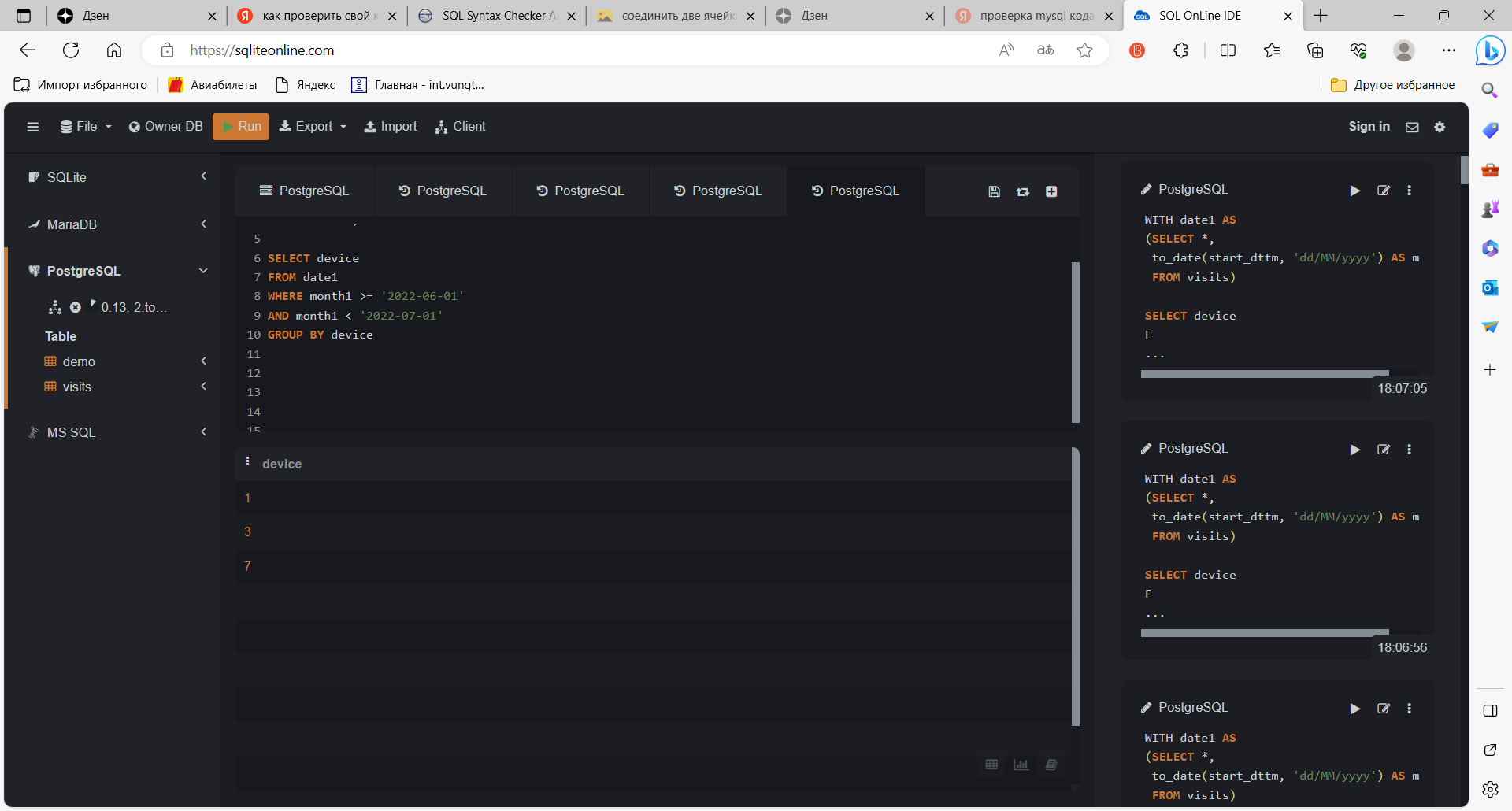
*--группировка по устройству*

*GROUP BY device*

*-- имеющему более 100 визитов*

*HAVING COUNT(\*) > 100;*

*В моей базе мало записей так что код в принт скрине выглядит без отбора в сотню*



***Задание 2.3***

***Выведите визиты всех пользователей, у которых никогда не было покупок.***

*-- таблицы слили по left join*

*SELECT*

*v.\*, p.uid*

*FROM*

*visits v*

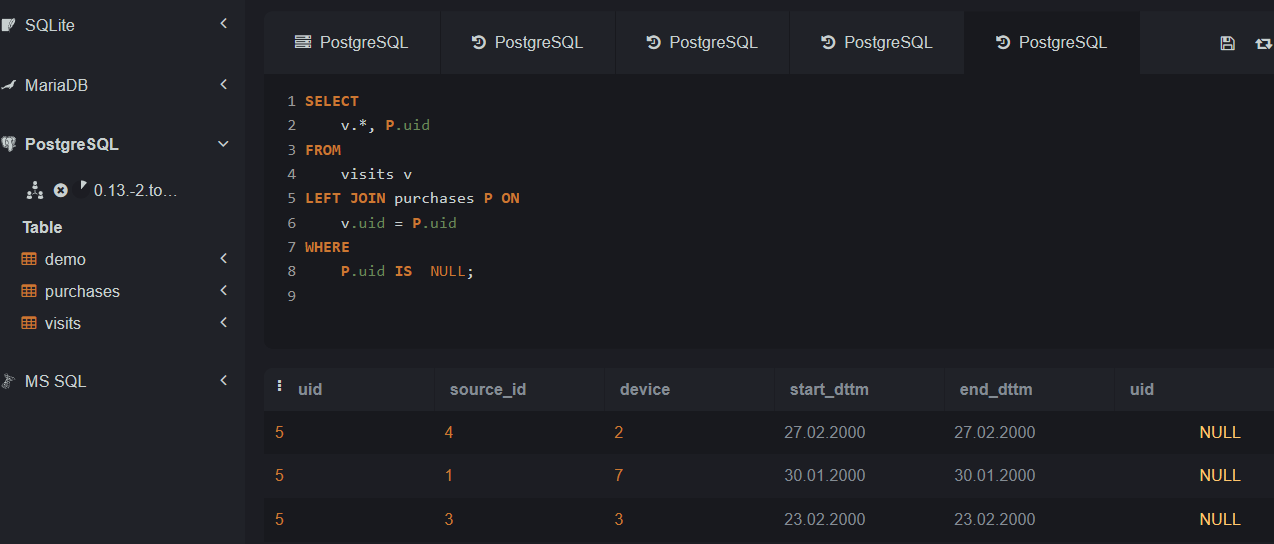
*LEFT JOIN purchases p ON*

*v.uid = p.uid*

*-- выводим визиты с нулом покупок*

*WHERE*

*p.uid IS NULL;*

**

***Задание 2.4***

***Для каждого клиента найдите средний чек за его последние три покупки.***

*SELECT p.uid, AVG(p.revenue) AS avg\_revenue*

*-- посчитали по колличеству номеров*

*FROM (*

*SELECT uid, revenue,*

*ROW\_NUMBER() OVER (PARTITION BY uid ORDER BY buy\_ts DESC) AS rn*

*FROM purchases*

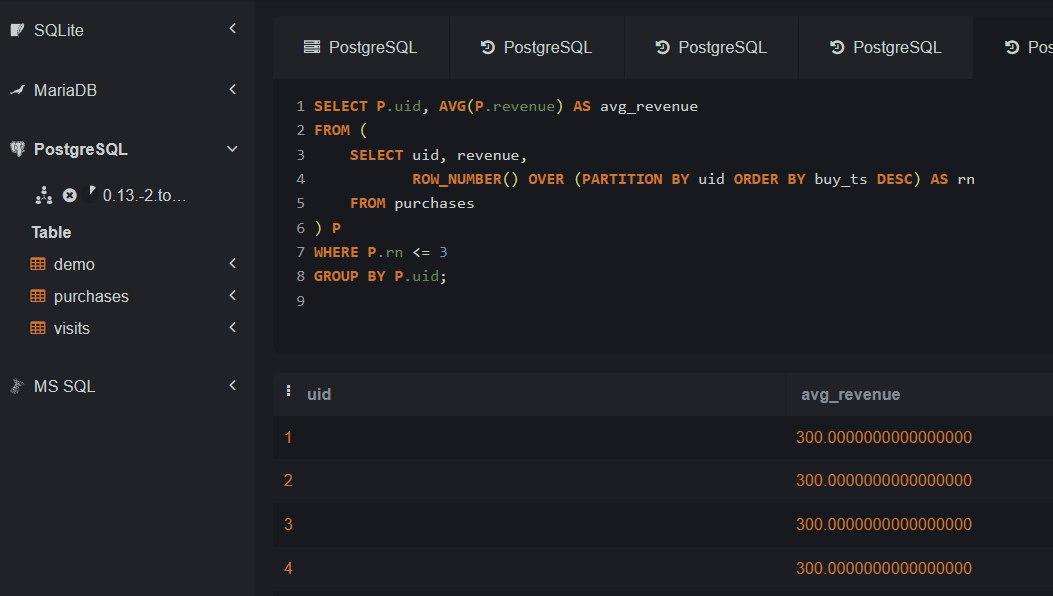
*) p*

*-- сделали отбор менее 3*

*WHERE p.rn <= 3*

*GROUP BY p.uid;*

*-- моя маленькая база как раз подходит*

**

### Задание 2.5

Для каждого источника визита рассчитайте среднее время, которое проходит между визитами. В расчет берите только визиты, продлившиеся дольше 10 секунд.

*SELECT*

*-- ОТБИРАЕМ ИСТОЧНИК*

*source\_id,*

*-среднее время между началом нового видита и выходом в прошлый визиь с группами по uid и по устройству в секундах*

*AVG(TIMESTAMP\_DIFF(start\_dttm, LAG(end\_dttm) OVER(PARTITION BY uid, source\_id ORDER BY start\_dttm), SECOND)) AS avg\_time\_between\_visits*

*FROM*

*visits*

*WHERE*

*-- где время визита более 10 секунд*

*TIMESTAMP\_DIFF(end\_dttm, start\_dttm, SECOND) > 10*

*GROUP BY*

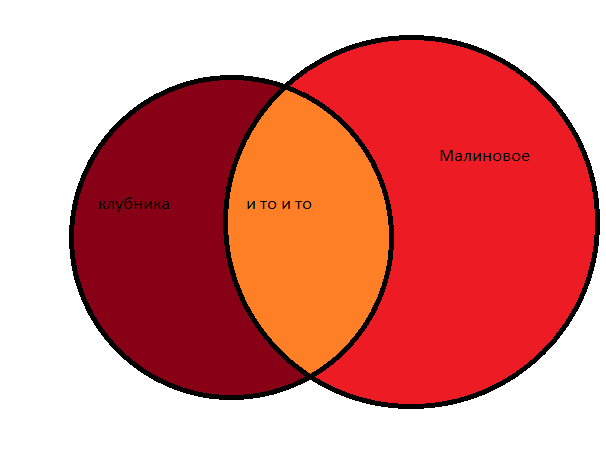
*source\_id*

## Задание 3

**Выберите верные, на ваш взгляд, ответы и укажите их в том же Word-файле, где вы сохранили решение по заданию 2.**

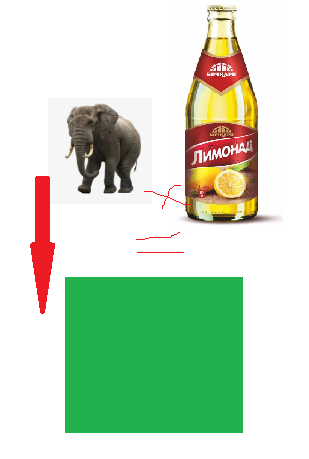
### Задание 3.1

Известно, что бабушки обязательно варят или малиновое варенье, или клубничное варенье. Могут варить и то и другое.



* Бабушка обязательно варит малиновое варенье. – нет, не обязательное условие
* Бабушка не может варить вишневое варенье вместо малинового и клубничного. Нет бабушки могут варить и вишневое, но варить или малиновое или клубничное варенье это обязательное условие.
* Бабушка обязательно варит и малиновое, и клубничное варенье. – бабушка может варить и клубничное и малиновое варенье, но может отдельно варить клубничное без малинового или малиновое без клубничного.

### Задача 3.2

**Если слон выпьет лимонад, то он станет зеленым.

* Если слон не зеленый, то он не пил лимонад. – верно. выпить лимонад достаточное условие быть зеленым
* Если слону не дать лимонад, то он не станет зеленым. не факт, он может облится краской. Зеленый не обязательно от лимонада
* Если слон зеленый, то он пил лимонад. – не факт, он мог облиться краской. Мог выпить и лимонад но быть зеленым не значит обязательно от лимонада

### Задача 3.4

Пальмы бывают или красными, или белыми. (Не (правда, что эта пальма не красная = белая) = красная).

* Эта пальма синяя.
* Эта пальма белая.
* Эта пальма красная.

### Задача 3.5

В природе обнаружено более десятка троллей. Все обнаруженные тролли зеленокожие. ( все известные троли зеленокожие, но могут быть не обнаруженные троли красных оттенков мы не можем на основании известного факта утверждать о единственном варианте)

* По крайней мере некоторые из троллей зеленокожие.
* Среди троллей есть краснокожие.
* Не существует не зеленокожих троллей.