# **База данных**

# **Задание 1**

У тебя есть база данных с поездками на такси. По плану на линию обслуживания должно было выйти 10550 автомобилей — эта цифра покрывает спрос пользователей. Команде поступило много жалоб: свободных автомобилей оказалось недостаточно. Сколько такси вышло на линии на самом деле? Информация о всех машинах на линии есть в таблице cabs.

Зайди на удалённый сервер.

Подключись к базе данных chicago\_taxi, используй логин morty и пароль smith.

Посчитай, сколько всего автомобилей в таблице cabs. Учти, что один автомобиль может принадлежать разным компаниями.

Число автомобилей:  
5500

Запрос, которым удалось решить задачу.  
SELECT COUNT(DISTINCT vehicle\_id) AS cnt\_vehicle FROM cabs;

## Задание 2

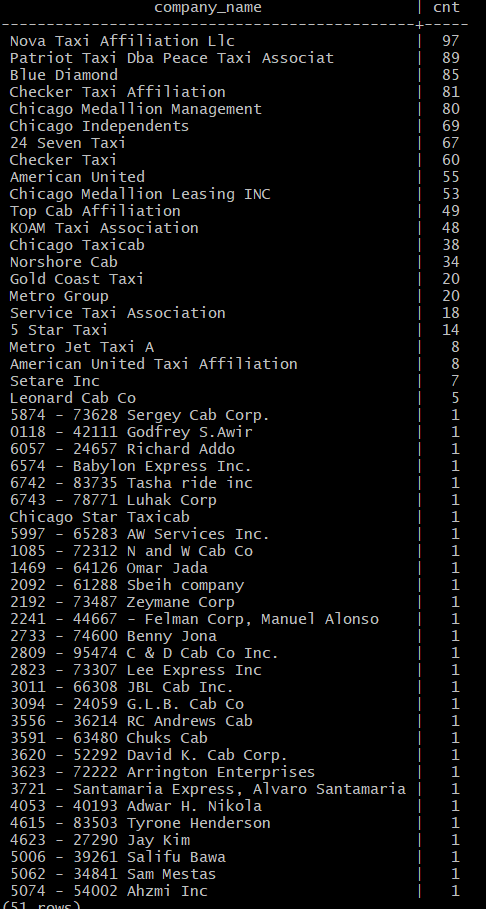
Посчитай количество автомобилей в каждой компании из таблицы cabs. Отсортируй значения по убыванию. Команда предполагает, что некоторые компании не вывели достаточно автомобилей на линию.

Выведи те компании, в которых меньше 100 автомобилей. Поле с числом автомобилей назови cnt, поле с названием компании — company\_name.

Чтобы решить задачу, примени оператор HAVING — аналог WHERE для агрегирующих функций. Изучи в документации, как работает оператор:

Список компаний с числом автомобилей меньше 100.

51 компания



Запрос, которым удалось решить задачу.

chicago\_taxi=# SELECT company\_name AS company\_name, COUNT(DISTINCT vehicle\_id) AS cnt FROM cabs GROUP BY company\_name HAVING COUNT(DISTINCT vehicle\_id)<100 ORDER BY cnt DESC;

## Задание 3

В приложении такси рассчитывается коэффициент стоимости поездки. Если погода хорошая, значение коэффициента равно 1. Если на улице дождь или шторм, коэффициент повышается до 2. У команды есть гипотеза, что в расчётах коэффициента ошибка. Чтобы проверить расчёт коэффициента, команде нужна выборка данных: разработчик может сверить коэффициент с данными в логах и исправить баг. Твоя задача — получить выборку.

Чтобы это сделать:

Получи описание погодных условий из таблицы weather\_records для каждого часа.

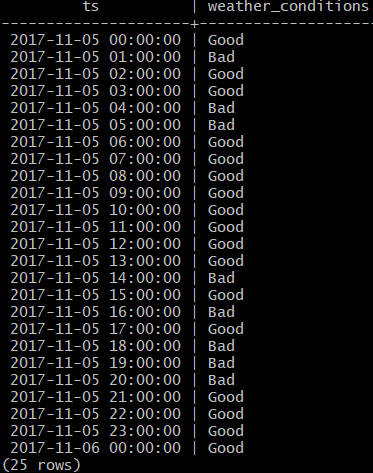
Раздели все часы на две группы оператором CASE: Bad, если поле description содержит слова rain или storm; Good для всех остальных.

Полученное поле назови weather\_conditions.

В результирующей таблице должно быть два поля — дата и час (ts) и weather\_conditions.

Сделай выборку за период с 2017-11-05 00:00 по 2017-11-06 00:00.

Таблица с данными за указанный период.



Запрос, которым удалось решить задачу.

SELECT ts, CASE WHEN description='rain' THEN 'Bad' WHEN description='storm' THEN 'Bad' ELSE 'Good' END AS weather\_conditions FROM weather\_records WHERE ts BETWEEN '2017-11-05 00:00' AND '2017-11-06 00:00';

ИСПРАВЛЕНО НА:

SELECT ts, CASE WHEN description LIKE '%rain%' THEN 'Bad' WHEN description LIKE '%storm%' THEN 'Bad' ELSE 'Good' END AS weather\_conditions FROM weather\_records WHERE ts BETWEEN '2017-11-05 00:00' AND '2017-11-06 00:00';

## Задание 4

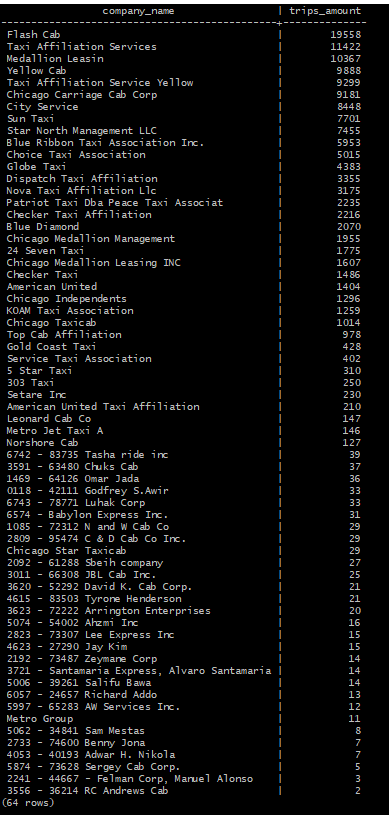
После обновления ПО таксопарки стали сообщать, что прибыль, которую они получают, не сходится с данными, которые отдаёт приложение. Разработка предполагает, что проблема может быть в данных о количестве поездок.

Чтобы определить, есть ли баг, нужно получить выборку с количеством поездок каждого таксопарка за 15 и 16 ноября 2017 года.

Выведи поле company\_name. Поле с числом поездок назови trips\_amount и выведи его.

Результаты, полученные в поле trips\_amount, отсортируй по убыванию.

Подсказка: чтобы решить задачу, соедини таблицы cabs и trips. Примени агрегирующие функции с группировкой. Не забудь написать конструкцию с условием.Таблица с данными за указанный период.



Запрос, которым удалось решить задачу.

chicago\_taxi=# SELECT DISTINCT cabs.company\_name AS company\_name, COUNT(DISTINCT trips.trip\_id) AS trips\_amount FROM cabs INNER JOIN trips ON trips.cab\_id=cabs.cab\_id WHERE trips.start\_ts BETWEEN '2017-11-15 00:00:00' AND '2017-11-16 23:59:59' GROUP BY company\_name ORDER BY trips\_amount DESC;