

avito.tech

# SQL И БАЗЫ ДАННЫХ

## SELECT. Выборка данных

# ПЛАН ЗАНЯТИЯ

## Повторение

Вспомним  
прошлые темы

## Теория+ Практика

Проходимся  
по темам из  
методички и  
решаем  
задачи

## Перерыв



10 минут

## Практика

Решаем задачи в  
парах

## ДЗ

Обсудим, что  
нужно будет  
сделать дома

1

Вот, что представляет аналитик, когда его попросили пофиксить баг в огромном SQL запросе, который писал не он



avito.tech

2

**Бэрримор, что это за жуткий вой в опенспейсе?**



Это аналитики. Они узнали, что сырых данных в БД нет, а есть только агрегаты.

3

# ПОВТОРЕНИЕ

Порядок выполнения  
логических операторов

**ORDER BY**

**Троичная  
логика**

**NULL**

**Предикат**

**Алиасы**

**DISTINCT**

# ОБСУЖДЕНИЕ ДЗ

01. `select model, speed, hd  
from pc where cd in ('12x', '24x') and price < 600`

02. `select item_id  
from item_descriptions  
where descr like '%//%'  
or descr like '%##%'`

03. `select name, EXTRACT(year from birthday) year  
from students  
where (name like '% A% %' and EXTRACT(year from birthday) < 1999)  
or (not name like '% % %' and name like '% A%' and EXTRACT(year from birthday) < 1999)  
or (name like '% П% %' and EXTRACT(year from birthday) > 1999)  
or (not name like '% % %' and name like '% П%' and EXTRACT(year from birthday) > 1999)`

# ОБСУЖДЕНИЕ ДЗ

04. `select distinct maker  
from product  
where type = 'PC'  
and maker not in (select maker  
from product  
where type = 'Laptop');`

05. `select city  
from cities  
order by city = 'Москва', city = 'Санкт-Петербург' desc, city desc;`

# ТЕОРИЯ

- ▶ Условные выражения
  - COALESCE
  - NULLIF
  - GREATEST, LEAST
  - CASE/CASE WHEN
- ▶ Агрегатные функции в простых выборках (без GROUP BY)
- ▶ COUNT
- ▶ SUM
- ▶ MIN / MAX / AVG
- ▶ Генерация последовательностей и случайных значений
- ▶ Использование DISTINCT в агрегациях
- ▶ GROUP BY
- ▶ Связь GROUP BY и ORDER BY
- ▶ GROUP BY по номерам столбцов



# Условные выражения COALESCE / NULLIF / GREATEST / LEAST CASE WHEN





Отсортируйте студентов из таблицы **students** по стипендии. В начало поставьте студентов со стипендией выше **20000**, остальных в конец. Внутри своей группы (ниже/выше **20000**) отсортируйте студентов в порядке возрастания стипендии.



# Агрегатные функции в простых выборках COUNT, SUM, MIN/MAX, AVG



Чем отличаются COUNT(\*) и  
COUNT(group\_code) ?



# Чем отличаются GREATEST и MAX ?



Напишите запрос к таблице **stocks**, который определит наибольшее количество NULL среди столбцов **open**, **high**, **low**.





По таблице **stocks** вычислите среднюю цену открытия. AVG мы еще не проходили =)



По таблице **stocks** найдите наибольший дневной рост цены акции, то есть разницы между ценой закрытия и открытия.



**DISTINCT в агрегациях. В каких агрегатных функциях нет смысла добавлять DISTINCT?**





Для чего используется GROUP BY?

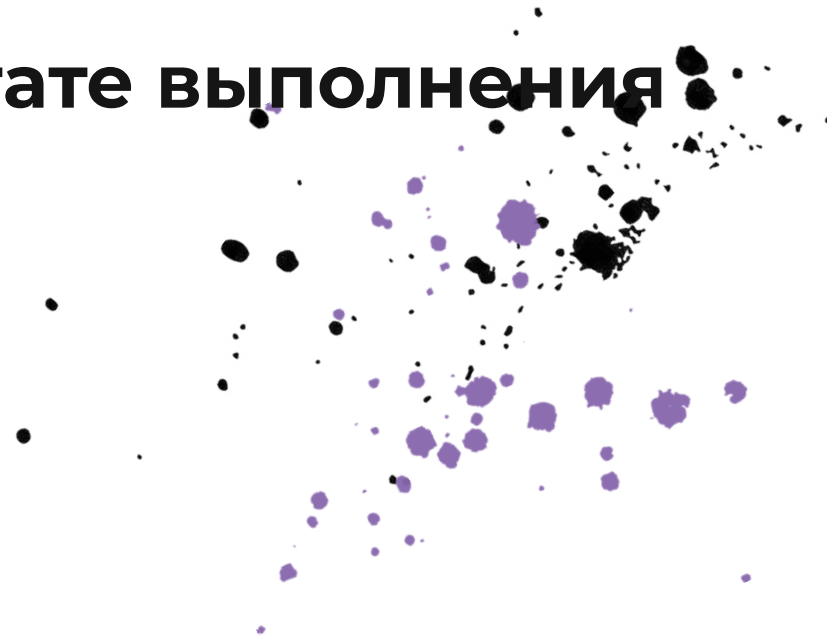


По таблице **stocks** найдите средний объем дневных торгов для каждого месяца. Результат отсортируйте в порядке убывания объема торгов.



# Что получится в результате выполнения запроса?

```
1 select mn,  
2     avg(volume) as avg_trade_volume  
3 from (  
4     select extract(month from date) as mn,  
5           volume,  
6           close  
7     from stocks  
8 ) stocks  
9 group by mn  
10 order by close desc;
```





# ПЕРЕРЫВ ВСТРЕЧАЕМСЯ ЧЕРЕЗ 10 МИНУТ





# Генерация последовательностей и случайных значений

```
SELECT * FROM generate_series(start,end);  
SELECT RANDOM();
```



Вычислите значение числа  $\pi$  численным  
методом Монте-Карло

<https://habr.com/ru/post/128454/>



# Вычислите значение числа $\pi$ численным методом Монте-Карло

$S_{\text{круга}} / S_{\text{квадрата}} = \text{Кол-во точек в круге} / \text{Кол-во точек в квадрате}$

$\pi \cdot (d/2)^2 / d^2 = \text{cnt\_in\_circle} / \text{cnt\_in\_square}$

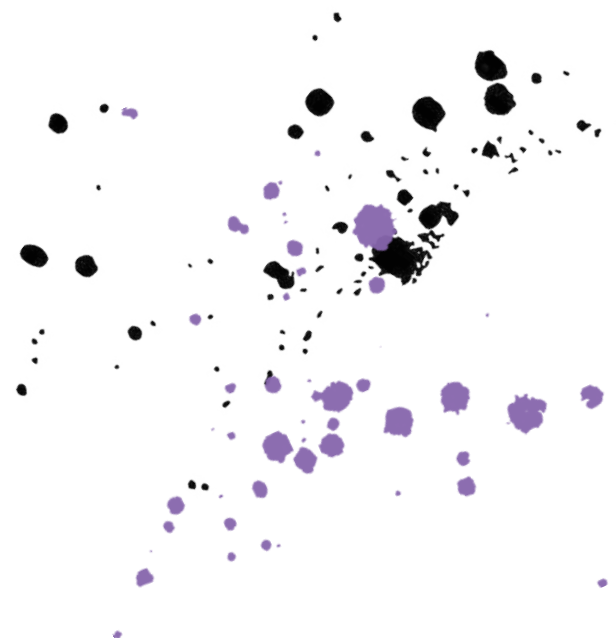
$\pi = 4 \cdot \text{cnt\_in\_circle} / \text{cnt\_in\_square}$

- 1) Задать достаточно много случайных точек на единичном квадрате
- 2) Проверить попадает ли точка в единичный круг внутри квадрата
- 3) Посчитать отношение попавших и не попавших точек

Можно совместить пункт 2 и 3



# ПРАКТИКА. Решаем задачи







По таблице **operations** найти для каждого опорного пункта (**base**) среднюю стоимость всех проводимых там операций.



По таблице **stocks** посчитайте объем торгов в каждом месяце. Выведите результат в хронологическом порядке.



Подсчитайте количество различных месяцев для каждого года по таблице **stocks**.



По таблице **PC** найдите модели компьютеров, имеющих самую высокую цену. Вывести: **model**, **price**.



Найдите среднюю скорость ПК, выпущенных производителем А. (Производителя можно узнать из таблицы Product)





По таблице **register** определите количество пятерок, четверок, троек у каждого студента.

# ДОМАШНЕЕ ЗАДАНИЕ



## 5 задач в stepik

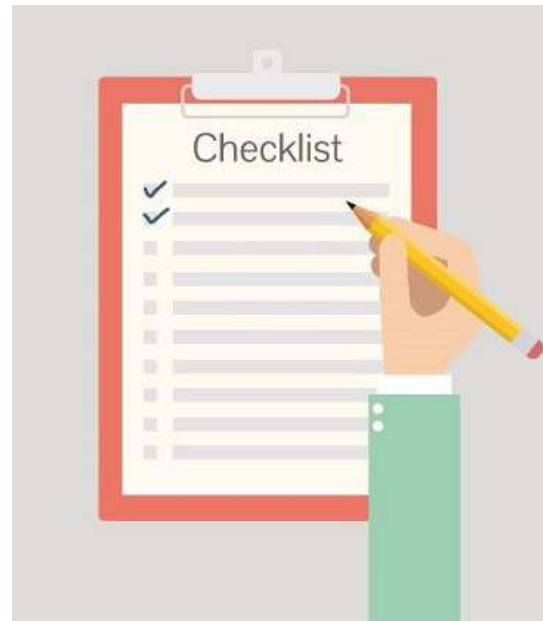
### Дедлайн:

ВС 10:00 - Мягкий (можно будет внести правки без потери баллов)

ПН 23:59 - Жесткий (баллы снижаются)

# ВЫВОДЫ

- ✓ познакомились с условными выражениями
- ✓ научились агрегировать данные
- ✓ узнали, как генерировать случайные последовательности







На следующем занятии продолжим с группировкой и научимся соединять таблицы.