avito.tech

SQL И БАЗЫ ДАННЫХ SELECT. Выборка данных

ПЛАН ЗАНЯТИЯ

Повторение РПрактика

Пройдемся по основным моментам из методички

Решаем задачи

Перерыв



10 минут

Практика

Решаем задачи

Д3

Обсудим, что нужно будет сделать дома

ПОВТОРЕНИЕ

Реляционная модель

СУБД

Связь

Сущность

Концептуальная модель

База

Отношение (таблица)

Кортеж

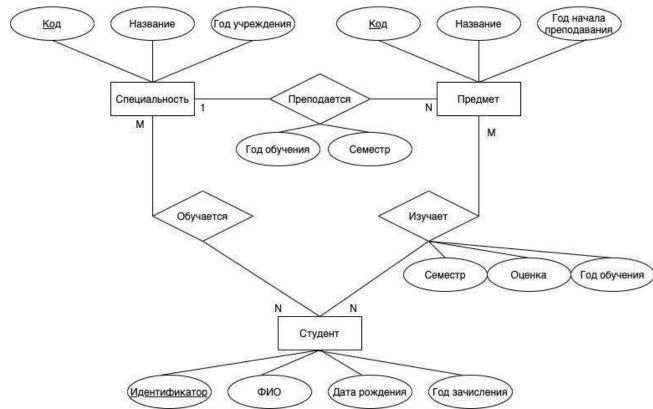
данных

Атрибут

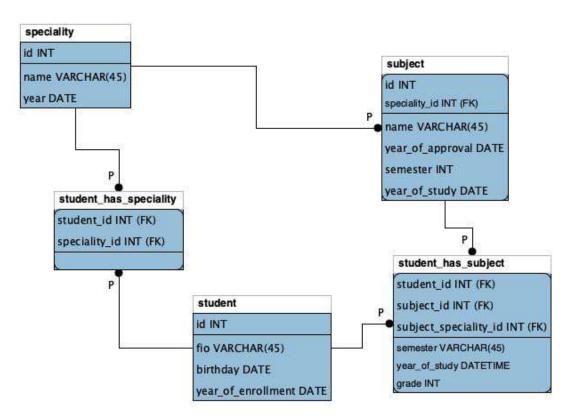
Внешний ключ

Проектирование

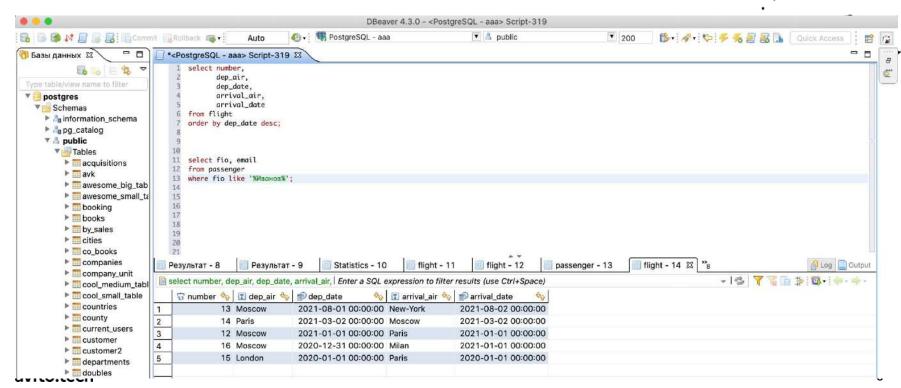
ОБСУЖДЕНИЕ ДЗ



ОБСУЖДЕНИЕ ДЗ







ТЕОРИЯ

- Оператор SELECT
- Выборка заданных столбцов
- Алиасы для столбцов
- Выражения в SELECTсписке
- Значение NULL
- DISTINCT
- Сортировка (ORDER BY)

- Фильтрация выборки (WHERE)
- Предикаты сравнения:=, >, <, >=, <=, <>
- Составные предикаты: AND, OR, NOT, скобки
- WHERE
- BETWEEN
- IN, IN с подзапросом
- > IS NULL



Для чего нужен оператор SELECT?





Как выбрать все атрибуты отношения?

Пример

```
presuser id.
pre.event_date,
pre is revenue.
isnull(case when (pre.revenue count daily = pre.revenue count cum) and (pre.revenue count daily >0) then True else False end, null) as is first from source,
pre first date from source,
pre.region_geo,
null as region_group,
pre.city_geo,
cu. IsCompany as user is company,
               when autype.AutoUploadType ext \Leftrightarrow 0 and cc.vertical = 'General' and ISNULL(user is subscriber.false) and stn.tariff detailed = 'subscr general gold' then true
               when autype.AutoUploadType ext <> 0 and cc.vertical = 'Jobs' then true
               else false
end user is autouploader.
ulp.user_id is not null as user_is_legal_entity,
cu.isBlocked as user is blocked,
uip.user id is not null as user personal report recipient,
uis.User_id is not null as user_is_shop,
usm.personal manager id,
usm.personal manager team,
CASE WHEN upm. start date is null THEN FALSE
                   WHEN upm.start_date=pre.event_date THEN day_has_shop_manager
                    ELSE day last has shop manager END AS user has pm,
case when pre-listing daily = pre-listing cum then True else False end as user lister isnew.
cu.RegistrationTime as user_registration_date,
cu.FirstListingDate as user_first_listing_date,
case when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event date) <= 3 then '1-3' when datediff('month', cu.RegistrationTime, pre.event date) <= 3 then '1-3' when datediff('month', cu.RegistrationTime, pre.event date) <= 3 then '1-3' when datediff('month', cu.RegistrationTime, pre.event date) <= 3 then '1-3' when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event date) <= 0 then '0' when datediff('month', cu.RegistrationTime, pre.event datediff('
case when revenue count cum <=0 then '0' when revenue count cum <=1 then '1' when revenue count cum <=3 then '2-3' when revenue count cum <=8 then '4-8' else '9+' end as sequence count cum <=0 then '1' when revenue count cum <=0 then '2-3' when revenue count cum <=0 then '1' when revenue cum <=0 the
case when listing cum <=0 then '0' when listing cum <=1 then '1' when listing cum <=3 then '2-3' when listing cum <=6 then '4-6' when listing cum <=12 then '7-12' when listing cum <=6 then '4-6' when listing cum <=12 then '7-12' when listing cum <=6 then '4-6' when listing cum <=12 then '7-12' when listing cum <=6 then '4-6' when listing cum <=12 then '7-12' when listing cum <=6 then '4-6' when listing cum <=12 then '7-12' when listing cum <=6 then '4-6' when listing cum <=12 then '7-12' when listing cum <=6 then '4-6' when listing cum <=12 then '7-12' when listing cum <=6 then '4-6' when listing cum <=12 then '7-12' when listing cum <=6 then '4-6' when listing cum <=12 then '7-12' when listing cum <=6 then '4-6' when listing cum <=12 then '7-12' when listing cum <=6 then '4-6' when listing cum <=6 then '4-6' when listing cum <=12 then '7-12' when listing cum <=6 then '4-6' when li
isnull(cc.ParentCategoryName, 'undefined') as category,
 isnull(cc.CategoryName, 'undefined') as subcategory,
cu.External_id as user_ext,
stn.tariff detailed,
```



Что такое NULL? Как получить все строки с NULL в нужном столбце?



Как выбрать уникальные значения?



Что сделает DISTINCT, если в исходной наборе данных несколько NULL?



Как отсортировать результат? • Какая сортировка по умолчанию?



Для чего используется WHERE?





Как определить диапазон значений?





Как проверять совпадение значений в столбце со значениями из списка? а из вложенного запроса?



Какой порядок выполнения операторов?

ORDER BY

WHERE

FROM

DISTINCT

SELECT



Какой запрос выдаст ошибку?

```
1 select distinct name, birthday
```

- 2 from students
- 3 order by student_id;

```
select name, birthday
```

- 2 from students
- 3 order by student_id;



Что получится в результате выполнения запроса?

```
1 SELECT
2    name,
3    extract (year from birthday) AS year_of_birth
4 FROM students
5 WHERE year_of_birth = 2000;
6
```



ПРАКТИКА. Решаем задачи





ПРО ТАБЛИЦУ students

select * from students 🚰 Введите SQL выражение чтобы отфильтровать результаты								
	123 student_id 🏋	RBC name T:	RDC group_code T:	② birthday T :	123 studentship Ҭ 🕻			
1	1	Иванов Иван Иванович	M8O-101A-19	2001-05-20	5 000			
2	2	Васильев Василий Васильевич	M8O-102A-19	2000-12-02	15 000			
3	3	Федоров Федор Федорович	M8O-103A-19	2001-01-17	25 000			
4	4	Григорьев Григорий Григорьевич	M8O-104A-19	2001-05-20	0			
5	5	Петров Петр Петрович	M8O-105A-19	2000-12-31	10 000			
6	6	Игнатьев Игнат Игнатьевич	M8O-102A-19	[NULL]	17 000			

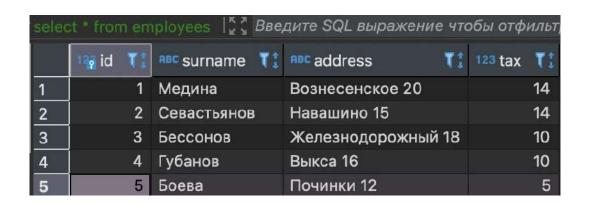


ПРО ТАБЛИЦУ delivery_orders

select * from delivery_orders 🎇 Введите SQL выражение чтобы отфильтровать рез							
	12 id T :	123 user_id 👯	123 products_count	TI	123 sum \(\frac{1}{3}\)	RBC status 👯	
1	1	1		2	1 300	new	
2	2	18		1	200	cancelled	
3	3	11		1	2 140	in_progress	
4	4	145		5	6 800	new	
5	5	23		1	999	new	
6	6	1		2	7 690	cancelled	



ПРО ТАБЛИЦУ employees





ПРО ТАБЛИЦУ operations

select $*$ from operations $ {}^{\kappa}_{\mathbf{x},\mathbf{y}} ^{\kappa}$ Введите SQL выражение чтобы отфильтровать результаты									
	¹≩id T ‡	RBC nomination T:	RBC base T:	123 reserve T‡	123 cost T ‡				
1	1	Наложение гипса	Выкса	2 000	18 000				
2	2	Блокада	Железнодорожный	10 000	14 000				
3	3	Инъекция поливитаминов	Навашино	20 000	11 000				
4	4	Инъекция алоэ	Железнодорожный	12 000	11 000				
5	5	ЭКГ	Вознесенское	115	10 000				
6	6	УЗИ	Вознесенское	20	30 000				



Проверьте по таблице students, является ли 2000 год годом рождения студента с id=5? id=6?



А СЕЙЧАС РАБОТА В ГРУППАХ :)





Выведите имя и группу студентов из таблицы students. При отсутствии группы выведите 'Unknown group'.



Напишите запрос к таблице студентов, который отсортирует студентов в порядке очередности празднования дней рождения.



Вывести различные группы, студенты которых получают стипендию не менее 20000.



ПЕРЕРЫВ ВСТРЕЧАЕМСЯ ЧЕРЕЗ 10 МИНУТ



Выберите из таблицы delivery_orders все заказы кроме отмененных, выведите всю информацию о них. У отмененных заказов статус 'cancelled'.



Напишите запрос к таблице студентов, который выбирает сначала студентов с четными ID, затем остальных. При этом в своей группе четности студенты отсортированы по убыванию ID.



Без использования оператора IN найдите студентов, обучающихся в группе М8О-101А-19 или М8О-104А-19 и имеющих id, кратный 3.



По таблицы employees определите размер налога для медперсонала из Выксы или Навашино.



Найдите по таблице operations название, стоимость и адрес опорного пункта для операций, в названии которых есть слово 'Инъекция', и стоящих более 10000 руб. Результат отсортируйте по адресу и стоимости.



Выберите из таблицы delivery_orders все отмененные заказы стоимостью от 3000 до 10000 рублей включительно.



Выберите из таблицы delivery_orders все отмененные (cancelled) или возвращенные (returned) заказы.



Выберите студентов из таблицы students, которые получают стипендию (studentship) меньше 30000 рублей и отсортируйте их по дате рождения. Студентов без группы выбирать не нужно.



Задача 13*

Отсортируйте таблицу cities таким образом, чтобы в начале находился Санкт-Петербург, в конце - Москва, а остальные города были отсортированы в лексикографическом порядке по убыванию.

ДОМАШНЕЕ ЗАДАНИЕ



4 задачи в stepik

Мягкий дедлайн: 25.09.2022 (ВС) 10:00

Жесткий дедлайн: 27.09.2022 (ВТ) 10:00

Требования к ДЗ

- 1. Код удовлетворяет стайл гайду (пришлем позднее)
- 2. Не хакнута система, т.е. автотест корректно сработал и код решения верный
- 3. В решении используется материал из лекций, запрос более-менее оптимальный

Подробнее - степик

(HO(HO(HONK)NK)NK)NK)



выводы

- начали знакомство с синтаксисом SQL
- узнали, как делать простые выборки данных
- написали свои первые запросы





На следующем занятии познакомимся с условными выражениями, агрегатными функциями и группировкой.