## **2. Предварительный анализ данных**

**Задание 2.1**

Рассчитайте максимальный возраст (max\_age) кандидата в таблице.

**Текст запроса:**

select

max(c.age)

from

hh.candidate c

**Результат:**



**Выводы:**

Максимальный возраст выглядит нереалистично.

### **Задание 2.2**

Теперь давайте рассчитаем минимальный возраст (min\_age) кандидата в таблице.

**Текст запроса:**

select

min(c.age)

from

hh.candidate c

**Результат:**



**Выводы:**

Минимальный возраст не вызывает никаких сомнений.

### **Задание 2.3**

Попробуем «почистить» данные. Напишите запрос, который позволит посчитать для каждого возраста (age) сколько (cnt) человек этого возраста у нас есть.

Отсортируйте результат по возрасту в обратном порядке.

**Текст запроса:**

select

c.age,

count(c.id) cnt

from

hh.candidate c

group by

c.age

order by

age desc

**Результат:**



**Выводы:**

Возраст 100 лет является очевидным выбросом выборки.

### **Задание 2.4**

По данным Росстата, средний возраст занятых в экономике России составляет 39.7 лет. Мы округлим это значение до 40. Найдите количество кандидатов, которые старше данного возраста. *Не забудьте отфильтровать «ошибочный» возраст 100.*

**Текст запроса:**

select

count(c.id) cnt

from

hh.candidate c

where

c.age>40 and c.age<>100

**Результат:**



**Выводы:**

Среди более чем 44 тысяч соискателей - примерно 14 % людей старше 40 лет.

## **3. Глобальный анализ показателей**

### **Задание 3.1**

Напишите запрос, который позволит узнать, сколько (cnt) у нас кандидатов из каждого города (city).

**Формат выборки:** city, cnt.

Группировку таблицы необходимо провести по столбцу title, результат отсортируйте по количеству в обратном порядке.

**Текст запроса:**

select

cit.title city,

count(can.id) cnt

from

hh.candidate can

join

hh.city cit on cit.id = can.city\_id

group by

city

order by

cnt desc

**Результат:**

****

### **Задание 3.2**

Москва бросается в глаза как, пожалуй, самый активный рынок труда. Напишите запрос, который позволит понять, каких кандидатов из Москвы устроит «проектная работа».

**Формат выборки:** gender, age, desirable\_occupation, city, employment\_type.

Отсортируйте результат по *id* кандидата.

**Текст запроса:**

select

gender,

age,

desirable\_occupation,

cit.title city,

employment\_type

from

hh.candidate can

join

hh.city cit on cit.id = can.city\_id

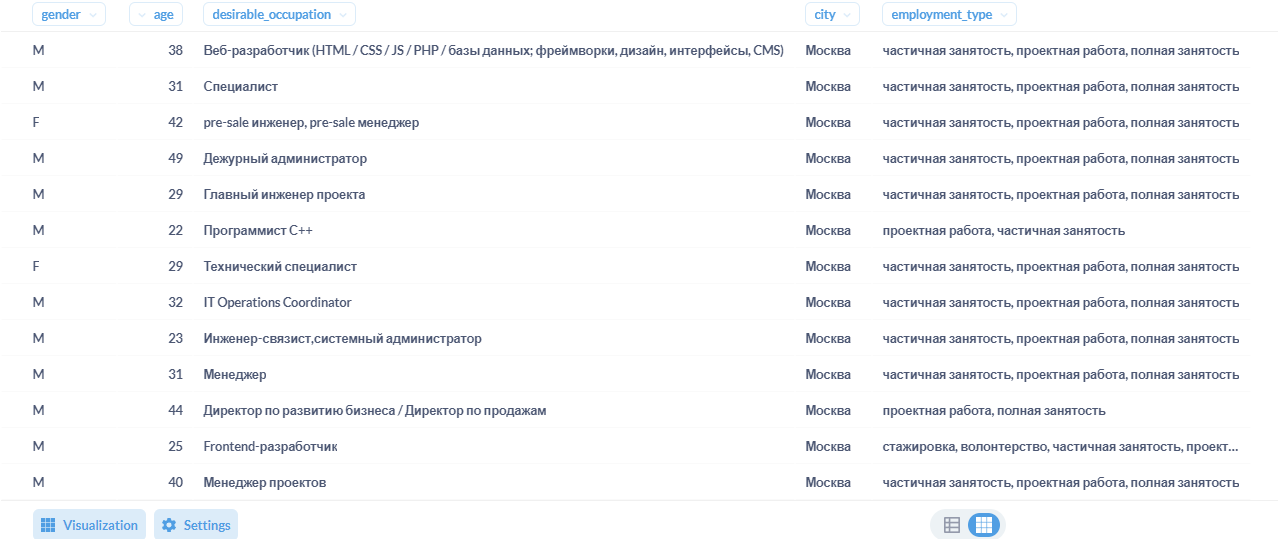
where

cit.title = 'Москва' and can.employment\_type like '%проектная работа%'

order by

can.id

**Результат:**

****

### **Задание 3.3**

Данных оказалось многовато. Отфильтруйте только самые популярные *IT*-профессии — разработчик, аналитик, программист.

Обратите внимание, что данные названия могут быть написаны как с большой, так и с маленькой буквы.

Отсортируйте результат по *id* кандидата.

**Текст запроса:**

select

gender,

age,

desirable\_occupation,

cit.title city,

employment\_type

from

hh.candidate can

join

hh.city cit on cit.id = can.city\_id

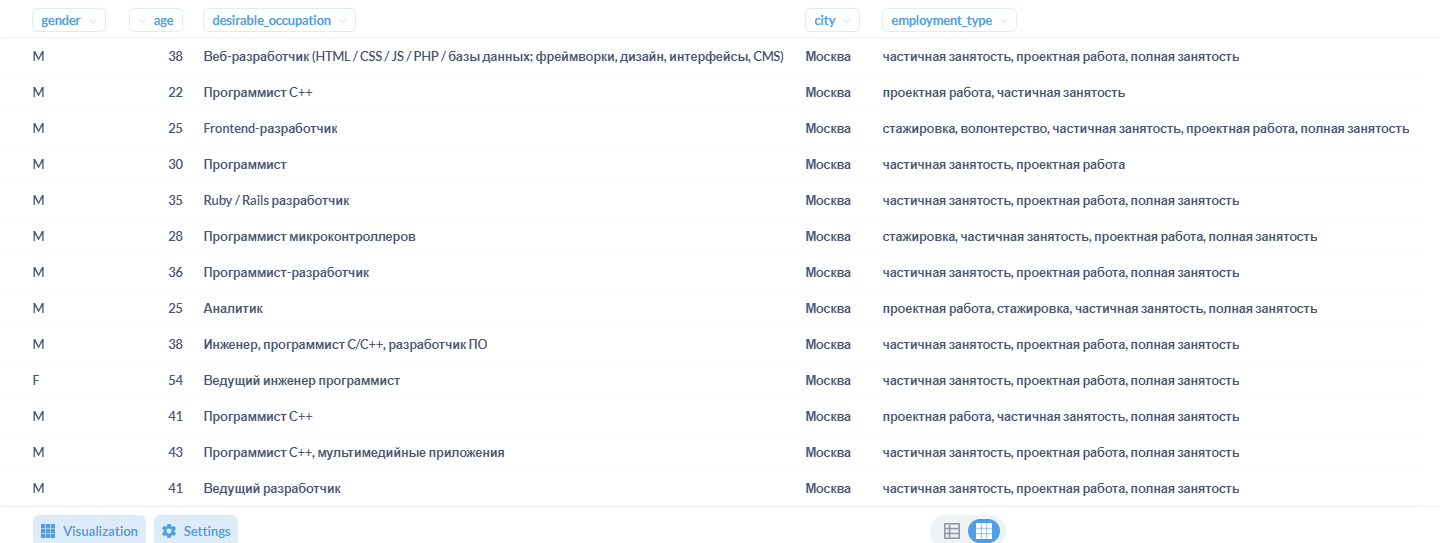
where

cit.title = 'Москва' and can.employment\_type like '%проектная работа%' and (desirable\_occupation ilike '%разработчик%' or desirable\_occupation ilike '%аналитик%' or desirable\_occupation ilike '%программист%')

order by

can.id

**Результат:**

****

### **Задание 3.4**

Для общей информации попробуйте выбрать номера и города кандидатов, у которых занимаемая должность совпадает с желаемой.

**Формат выборки:** id, city.

Отсортируйте результат по городу и *id* кандидата.

**Текст запроса:**

select

can.id,

cit.title city

from

hh.candidate can

join

hh.city cit on cit.id = can.city\_id

where

desirable\_occupation = current\_occupation

order by

city,

can.id

**Результат:**

****

### **Задание 3.5**

Определите количество кандидатов пенсионного возраста.

*Пенсионный возраст для мужчин наступает в 65 лет, для женщин — в 60 лет.*

**Текст запроса:**

select

count(id)

from

hh.candidate

where

((gender = 'M' and age>=65) or (gender = 'F' and age>=60)) and age<>100

**Результат:**

****

## **4. Анализ кандидатов для заказчиков**

### **Задание 4.1**

Для добывающей компании нам необходимо подобрать кандидатов из Новосибирска, Омска, Томска и Тюмени, которые готовы работать вахтовым методом.

**Формат выборки:** gender, age, desirable\_occupation, city, employment\_type, timetable\_type.

Отсортируйте результат по городу и номеру кандидата.

**Текст запроса:**

select

gender,

age,

desirable\_occupation,

cit.title city,

employment\_type,

tt.title timetable\_type

from

hh.candidate can

left join

hh.city cit on cit.id = can.city\_id

left join

hh.candidate\_timetable\_type ctt on ctt.candidate\_id = can.id

left join

hh.timetable\_type tt on tt.id = ctt.timetable\_id

where

(cit.title='Омск' or

cit.title='Томск' or

cit.title='Тюмень' or

cit.title='Новосибирск') and

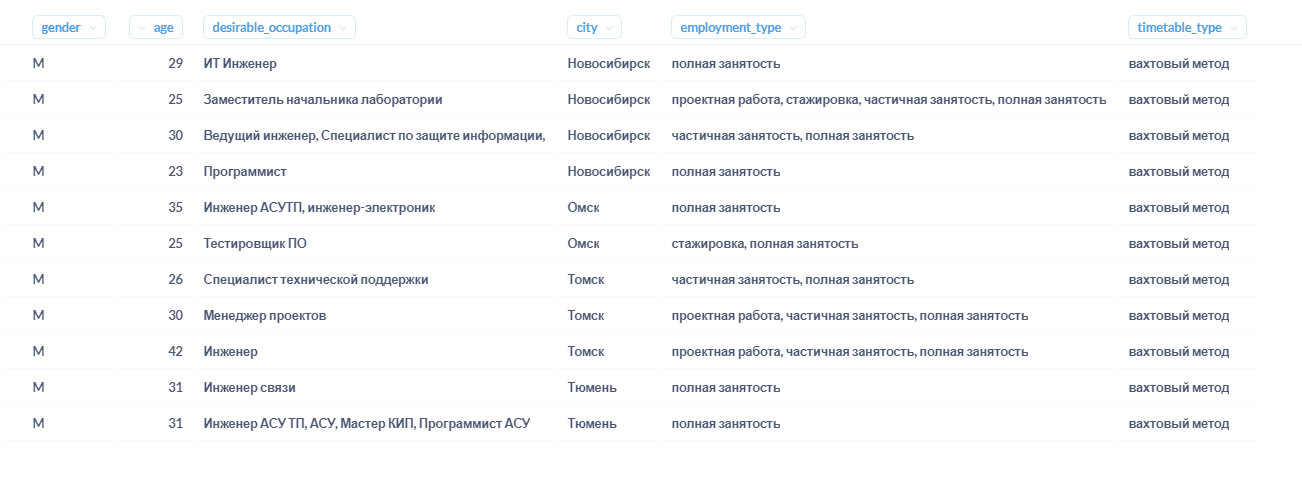
tt.id=4

order by

city,

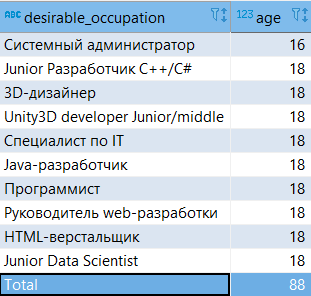
can.id

**Результат:**

****

### **Задание 4.2**

Для заказчиков из Санкт-Петербурга нам необходимо собрать список из 10 желаемых профессий кандидатов из того же города от 16 до 21 года (в выборку включается 16 и 21, сортировка производится по возрасту) с указанием их возраста, а также добавить строку Total с общим количеством таких кандидатов. Напишите запрос, который позволит получить выборку вида:



**Текст запроса:**

(select

desirable\_occupation,

age

from

hh.candidate can

left join

hh.city cit on cit.id = can.city\_id

where

cit.title='Санкт-Петербург' and age between 16 and 21

order by

2

limit

10)

union all

(select

'Total',

count(can.id)

from

hh.candidate can

left join

hh.city cit on cit.id = can.city\_id

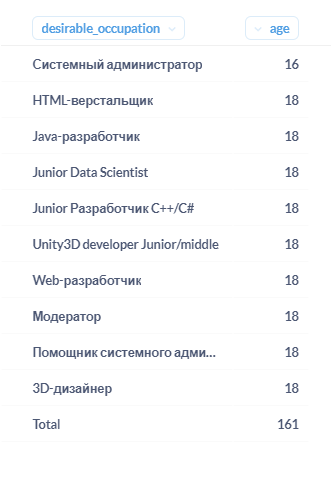
where

cit.title='Санкт-Петербург' and age between 16 and 21)

order by

2

**Результат:**

****