

PROJECT-2. Подгрузка новых данных. Уточнение анализа

Выполнил Борзаковский Владислав Евгеньевич

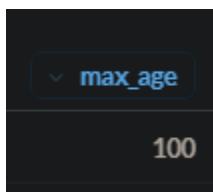
Задание 2.1

Рассчитайте максимальный возраст (`max_age`) кандидата в таблице.

Текст запроса:

```
SELECT max(age) max_age /*при помощи функции max определяем
максимальный возраст*/
FROM hh.candidate
```

Результат:



max_age
100

Выводы:

Максимальный возраст кандидата 100 лет. Данный результат вероятно какая-то ошибка, случайный ввод или кандидат намеренно скрыл свой настоящий возраст.

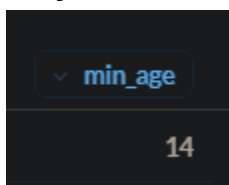
Задание 2.2

Теперь давайте рассчитаем минимальный возраст (`min_age`) кандидата в таблице.

Текст запроса:

```
SELECT min(age) min_age /*при помощи функции min определяем минимальный
возраст*/
FROM hh.candidate
```

Результат:



min_age
14

Выводы:

Минимальный возраст кандидата 14-ть лет. Данный результат похож на истину.

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

age, count(age) cnt /*выбираем столбец age и при помощи count подсчитываем количество кандидатов*/

from hh.candidate

group by 1 /*группируем по первому столбцу*/

order by 1 desc /*сортируем по возрасту в обратном порядке*/

Результат:

age	cnt
100	1
77	1
76	1
73	4
72	3
71	4

Выводы:

Кандидаты встречаются разных возрастов и по несколько раз. Возраст 100 встречается один раз, что ещё раз подтверждает не достоверность этого значения. В дальнейших анализах этот возраст лучше не учитывать.

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(*) /*вызываем счетчик*/

from hh.candidate

where age > 40 and age < 100 /*вводим условия фильтра*/

Результат:

Количество
6,263

Выводы:

Мы нашли количество кандидатов старше 40-ка лет. Оно равняется 6263-ём. Это примерно седьмая часть от всех кандидатов.

Задание 3.1

Для начала 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

title city, count(candidate.id) cnt /*при помощи счётчика считаем количество кандидатов в каждом городе*/

from hh.candidate as candidate

join hh.city as city on city.id = candidate.city_id /*присоединяем таблицу городов*/

group by title /*группируем по столбцу title*/

order by 2 desc /*сортируем по количеству в обратном порядке*/

Результат:

city	cnt
Москва	16,622
Санкт-Петербург	4,937
Краснодар	1,066
Новосибирск	958
Казань	872
Екатеринбург	734

Выводы:

Мы узнали сколько кандидатов в каждом городе. Преимущественно выигрывает город Москва. Стоит предположить, что это из-за количества населения.

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
  candidate.gender,
  candidate.age,
  candidate.desirable_occupation,
  city.title as city,
  candidate.employment_type /*формат выборки согласно заданию*/
from hh.candidate as candidate /*выбираем таблицу кандидатов*/
join hh.city as city on city.id = candidate.city_id /*присоединяем таблицу городов*/
where city.title = 'Москва' and candidate.employment_type like '%проектная
работа%' /*вводим условия фильтра для выполнения задания*/
order by candidate.id /*сортируем по ID кандидата*/
```

Результат:

gender	age	desirable_occupation	city	employment_type
М	38	Веб-разработчик (HTML / CSS / JS / PHP / базы данных; фреймворки, дизайн, интерфейсы, CMS)	Москва	частичная занятость, проектная работа, полная занятость
М	31	Специалист	Москва	частичная занятость, проектная работа, полная занятость
Ф	42	pre-sale инженер, pre-sale менеджер	Москва	частичная занятость, проектная работа, полная занятость
М	49	Дежурный администратор	Москва	частичная занятость, проектная работа, полная занятость
М	29	Главный инженер проекта	Москва	частичная занятость, проектная работа, полная занятость
М	22	Программист C++	Москва	проектная работа, частичная занятость

Выводы:

Мы видим, что проектная работа подходит в основном для кандидатов из Москвы, которые причастны к IT-сфере.

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
    candidate.gender,
    candidate.age,
    candidate.desirable_occupation,
    city.title as city,
    candidate.employment_type /*формат выборки согласно заданию*/
from hh.candidate as candidate /*выбираем таблицу кандидатов*/
join hh.city as city on city.id = candidate.city_id /*присоединяем таблицу городов*/
where
    city.title = 'Москва' and candidate.employment_type like '%проектная
    работа%' and
    (candidate.desirable_occupation ilike '%разработчик%' or
    candidate.desirable_occupation ilike '%программист%' or
    candidate.desirable_occupation ilike '%аналитик%') /*вводим условия фильтра
    для выполнения задания*/
order by candidate.id /*сортируем по ID кандидата*/
```

Результат:

gender	age	desirable_occupation	city	employment_type
М	38	Веб-разработчик (HTML / CSS / JS / PHP / базы данных; фреймворки, дизайн, интерфейсы, CMS)	Москва	частичная занятость, проектная работа, полная занятость
М	22	Программист C++	Москва	проектная работа, частичная занятость
М	25	Frontend-разработчик	Москва	стажировка, волонтерство, частичная занятость, проект...
М	30	Программист	Москва	частичная занятость, проектная работа
М	35	Ruby / Rails разработчик	Москва	частичная занятость, проектная работа, полная занятость
М	28	Программист микроконтроллеров	Москва	стажировка, частичная занятость, проектная работа, вол...

Выводы:

Мы отфильтровали только самые популярные *IT*-профессии. И видим, что среди кандидатов есть программисты разных языков, разработчики разных ПО. Так же есть кандидаты, у которых указано несколько профессий.

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
```

```

cnd.id,
ct.title city /*формат выборки согласно заданию*/
from hh.candidate cnd /*выбираем таблицу кандидатов*/
join hh.city ct on cnd.city_id = ct.id /*присоединяем таблицу городов*/
where cnd.desirable_occupation = cnd.current_occupation /*вводим условия
фильтра для выполнения задания*/
group by ct.title, cnd.id /*группируем по городу и id кандидата*/

```

Результат:

id	city
2,009	Абакан
10,340	Абакан
14,449	Абакан
20,261	Абакан
13,705	Агрыз
967	Адлер

Выводы:

Для общей информации мы узнали номера и города кандидатов, у которых занимаемая должность совпадает с желаемой. Количество таких кандидатов небольшое. Соответственно можно предположить, что люди именно по этому и хотят сменить работу.

Задание 3.5

Определите количество кандидатов пенсионного возраста.

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

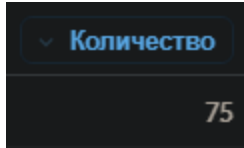
Текст запроса:

```

select
count(*) /*при помощи счётчика определяем количество*/
from hh.candidate /*выбираем таблицу кандидатов*/
where (gender = 'M' AND age >= 65 AND age < 100)
OR (gender = 'F' AND age >= 60 AND age < 100) /*вводим условия фильтра для
выполнения задания, не забываем о недостоверном значении возраста*/

```

Результат:



Выводы:

Количество кандидатов пенсионного возраста 75. Это очень малый процент от общего количества кандидатов. А это значит, что кандидаты пенсионного возраста не нуждаются в поиске работы.

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
    cd.gender gender,
    cd.age age,
    cd.desirable_occupation desirable_occupation,
    ct.title city,
    cd.employment_type employment_type,
    tt.title timetable_type /*формат выборки согласно заданию*/
FROM hh.candidate cd /*выбираем таблицу кандидатов*/
    JOIN hh.city ct ON cd.city_id = ct.id /*присоединяем таблицу городов*/
    JOIN hh.candidate_timetable_type ctt ON cd.id = ctt.candidate_id
/*присоединяем таблицу подходящих графиков*/
    JOIN hh.timetable_type tt ON ctt.timetable_id = tt.id /*присоединяем таблицу
рабочих графиков*/
WHERE (ct.title IN ('Новосибирск', 'Омск', 'Томск', 'Тюмень')) AND (tt.title
='вахтовый метод') /*вводим условия фильтра для выполнения задания*/
ORDER BY 4, cd.id /*сортируем результат по городу и номеру кандидата*/
```

Результат:

gender	age	desirable_occupation	city	employment_type	timetable_type
M	29	ИТ Инженер	Новосибирск	полная занятость	вахтовый метод
M	25	Заместитель начальника лаборатории	Новосибирск	проектная работа, стажировка, частичная занятость, полная занятость	вахтовый метод
M	30	Ведущий инженер, Специалист по защите информации,	Новосибирск	частичная занятость, полная занятость	вахтовый метод
M	23	Программист	Новосибирск	полная занятость	вахтовый метод
M	35	Инженер АСУТП, инженер-электроник	Омск	полная занятость	вахтовый метод
M	25	Тестировщик ПО	Омск	стажировка, полная занятость	вахтовый метод

Выводы:

Мы выбрали кандидатов. Их не много. Все мужчины, возраст от 25 до 42. Должность преимущественно “Инженер”.

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:

desirable_occupation	age
Системный администратор	16
Junior Разработчик C++/C#	18
3D-дизайнер	18
Unity3D developer Junior/middle	18
Специалист по IT	18
Java-разработчик	18
Программист	18
Руководитель web-разработки	18
HTML-верстальщик	18
Junior Data Scientist	18
Total	88

Текст запроса:


```

(select
  cnd.desirable_occupation,
  cnd.age /*формат выборки согласно заданию*/
from hh.candidate as cnd /*выбираем таблицу кандидатов*/
  join hh.city on cnd.city_id=city.id /*присоединяем таблицу городов*/
where
  city.title = 'Санкт-Петербург'
  and cnd.age between 16 and 21 /*вводим условия фильтра для выполнения
задания*/
order by cnd.age /*сортируем по возрасту*/
limit 10 /*делаем ограничение на 10 профессий*/)

union all /*оператор присоединения*/

(select /*добавляем строку Total*/
  'Total', count(*) /*при помощи счётчика произведём подсчёт*/
from hh.candidate as cnd /*выбираем таблицу кандидатов*/
  join hh.city on cnd.city_id=city.id /*присоединяем таблицу городов*/
where
  city.title = 'Санкт-Петербург'
  and age between 16 and 21 /*вводим условия фильтра для выполнения
задания*/)

```

Результат:

desirable_occupation ▾	age ▾
Системный администратор	16
Junior Разработчик C++/C#	18
Программист	18
Junior Data Scientist	18
Руководитель web-разработки	18
Специалист по IT	18
Unity3D developer Junior/middle	18
HTML-верстальщик	18
3D-дизайнер	18
Java-разработчик	18
Total	161

Выводы:

Из полученных результатов мы видим, что список состоит только из IT-профессий. Всего кандидатов 161. Ребята уже на школьном образовании готовы покорять данную сферу.

Общий вывод по проекту:

И так, проведя анализ мы узнали следующее:

- возраст кандидатов от 14 и до 77;
- так же есть процент людей старше 40-ка лет и даже пенсионного возраста;
- самый большой рынок труда в Москве, Петербург на втором месте, Краснодар на третьем;
- рынок труда зависит от плотности и количества населения;
- сфера IT-профессий самая популярная, а именно разработчик, аналитик, программист;
- так же есть кандидаты, у которых желаемая и действительная должности совпадают, но таких не много.

Исходя из всего этого можно понять, что люди меняют работу либо для смены сферы деятельности (после института поработали и не понравилось), либо для собственного роста со сменой должности. Молодые ребята не имея высшего образования уже ищут себе работу в сфере IT.

Мой вывод таков, что за IT будущее и надо быстрее в него вливаться.

