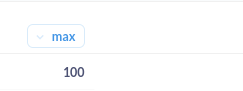
Задание 2.1

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

select 

max(can.age)

from hh.candidate can

*Это явный выброс значения в поле возраст. Кандидат скорее всего “шутник” или произошла ошибка при вводе значения. При дальнейшем анализе датасета, необходимо будет избавиться от подобный значений.*

Задание 2.2

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

select

min(can.age)

from hh.candidate can

*Значение также находится вне диапазона. Легально в РФ человек может трудоустроиться по достижении 16 лет. Все записи с возрастом менее 16 лет можно не включать в выборки для анализа.*

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.

select

can.age age,

count(\*) cnt

from hh.candidate can

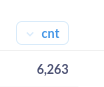
group by can.age

order by age desc

*Из полученных данных следует, что распределение кандидатов по возрасту носит нормальное (или ближе к логнормальному) распределение. Основная часть резюме сконцентрировано в диапазоне от 20 до 45 лет с максимумом распределения в 30 лет. Т.е. как и следовало ожидать, наиболее активны в поиске работы люди начиная с возраста студентов последних курсов или выпускников средних специальных учебных заведений и до возраста основной деловой активности 35-40 лет, после чего уже количество кандидатов резко уменьшается, так как либо карьера уже состоялась либо есть сложности с тем, что “возрастных” кандидатов на нашем рынке труда не очень охотно берут(по крайней мере есть такой миф).*

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.*

select 

count(\*) cnt

from hh.candidate can

where can.age > 40 and can.age < 100

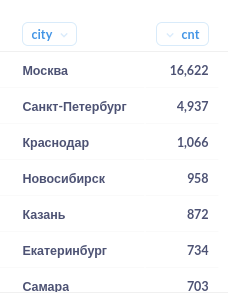
*Резюме, с указанным возрастом более 40 лет, составляют 14% от общего количества. Здесь можно рассуждать разными способами:*

*Во-первых необходимо уточнение термина “занятый в экономике”. Если это человек который трудоустроен и не ищет работу, то возможно он не имеет резюме на hh.ru и тогде картина выглядит реалистично: после 40 лет как раз меньшая часть резюме.*

*Во-вторых, если принять за истину, что всю занятость в экономике можно оценивать по нашей выборке на основе резюме на hh.ru (что мне представляется не совсем верным), то тогда данные Росстата немного “лукавят”, так как в данном случае идет подмена медианного возраста средним.*

Задание 3.1

Для начала напишите запрос, который позволит узнать, сколько (cnt) у нас кандидатов из каждого города (city). Отсортируйте результат по количеству в обратном порядке.

select 

c.title city,

count(\*) cnt

from hh.candidate can

join hh.city c on can.city\_id = c.id

group by c.title

order by cnt desc

*Вполне логично что наиболее активно заняты поиском работы кандидаты в Москве (более 37%). Также основные города миллионники на первых позициях в этом рейтинге. Это города имеющие наиболее развитую инфраструктуру, промышленность и создающие в стране наибольшее количество рабочих мест как на крупных предприятиях так и в сфере мелкого и среднего бизнеса.*

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.

select

can.gender,

can.age,

can.desirable\_occupation,

c.title city,

can.employment\_type

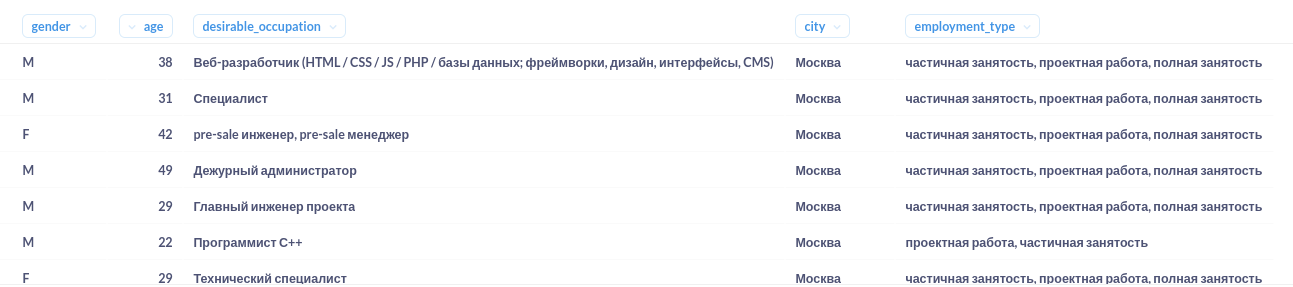
from hh.candidate can

join hh.city c on can.city\_id = c.id

where c.title = 'Москва' and -- city filtering

can.employment\_type like '%проектная работа%' -- type filtering for selection

order by can.id



*На проектную работу готовы по большей части мужчины в возрасте от 20 до 50 лет, следующих основных специальностей: программисты и разработчики, инженерных специальностей, менеджеры проектов или SMM-менеджеры, а также специалисты по системам бухучета типа 1С SAP и подобным.*

*Очевидно что не во всех отраслях и специализациях есть возможность работать с режимом проектной занятости. Однако существуют и наоборот такие типы специальностей, которые предполагают именно занятость в конкретных проектах на конкретный временной интервал, без привязки к штатному расписанию.*

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.

select

can.gender,

can.age,

can.desirable\_occupation,

c.title city,

can.employment\_type

from hh.candidate can

join hh.city c on can.city\_id = c.id

where c.title = 'Москва' and

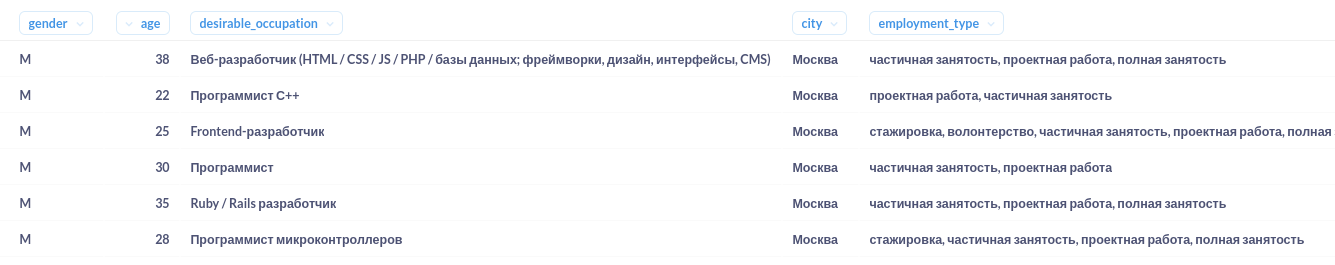
can.employment\_type like'%проектная работа%' and

(lower(can.desirable\_occupation) like ('%разработчик%') or -- check if sample1 include

lower(can.desirable\_occupation) like ('%аналитик%') or -- check if sample2 include

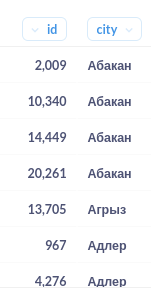
lower(can.desirable\_occupation) like ('%программист%')) -- check if sample3 include

order by can.id



*Данная выборка подтверждает предположение о том, что в основном мужчины в возрасте от 20 до 45 лет рассматривают варианты проектной работы. Все типы желаемой должности совместимы с режимом проектной работы и в основном это IT или инженерные позиции.*

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.

select

can.id,

c.title city

from hh.candidate can

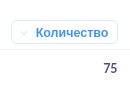
join hh.city c on can.city\_id = c.id

where can.desirable\_occupation = can.current\_occupation

order by city, can.id

*Оказалось, достаточное частое явление, когда соискатель рассматривает идентичную занимаемой должности, но видимо хочет увеличить зарплату или сменить компанию и прочие условия. Это также может быть связано с тем, что не все люди четко понимают, что именно бы им хотелось изменить в карьере и указывают такую же должность, так как в том что они ей соответствуют, они хотя бы уверены.*

Задание 3.5

Определите количество кандидатов пенсионного возраста. *Пенсионный возраст для мужчин наступает в 65 лет, для женщин — в 60 лет.*

select

count(\*)

from hh.candidate can

where

(can.gender = 'M' and can.age between 65 and 99) or -- filtering retired men

(can.gender = 'F' and can.age between 60 and 99) -- filtering retired women

*Хоть и совсем незначительная часть (сотые доли процента), но желающие трудоустроиться пенсионеры имеются. В реальности процент занятых в экономике пенсионеров я полагаю значительно больше. Пожилым людям составляет вполне определенную сложность разместить свое резюме на подобных hh.ru ресурсах (в 2022 году выходят на пенсию люди 1957, 1962 гг. рождения). В дальнейшее время, когда придет возраст выходить на пенсию следующим поколениям, которые уже застали бурный рост компьютерных технологий и интернет в более молодом возрасте, доля резюме пенсионеров на hh.ru и подобных агрегаторах только резко увеличится. В нашей стране, по крайнем мере в настоящее время и ближайшем обозримом будущем, традиционно люди продолжают трудиться на пенсии, если позволяет здоровье и условия найма.*

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.

select

can.gender,

can.age,

can.desirable\_occupation,

c.title city,

can.employment\_type,

tt.title timetable\_type

from hh.candidate can

join hh.city c on c.id = can.city\_id -- joining

join hh.candidate\_timetable\_type ctt on can.id = ctt.candidate\_id -- for timetable

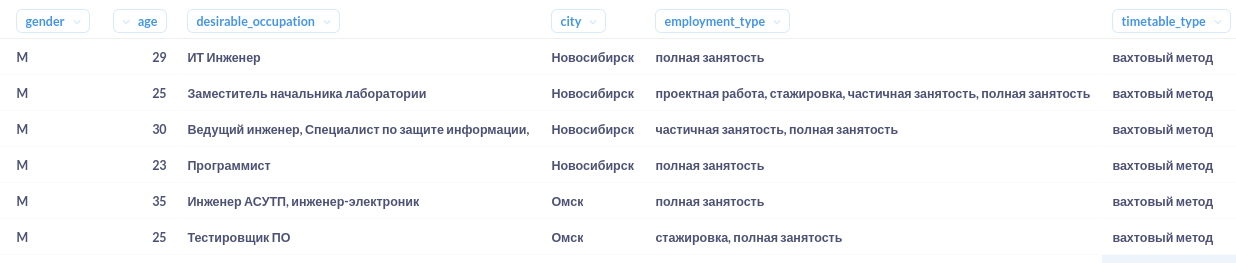
join hh.timetable\_type tt on ctt.timetable\_id = tt.id -- extraction

where

c.title in ('Новосибирск','Омск','Томск','Тюмень') and

lower(tt.title) ='вахтовый метод'

order by c.title, can.id



*Данная выборка показала низкую готовность работать вахтовым методом: 11 мужчин в возрасте от 23 до 42 лет, в основном технические и инженерные специальности. Вахтовый метод подразумевает долгосрочные командировки по расписанию, что является весомым ограничением для семейных мужчин, даже если уровень зарплат выше чем на постоянном месте работы.*

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:

(select

can.desirable\_occupation,

can.age

from hh.candidate can

join hh.city c on c.id = can.city\_id

where

c.title = 'Санкт-Петербург' and

can.age between 16 and 21

order by 2

limit 10)

union all

select

'Total',

count(\*)

from hh.candidate can

join hh.city c on c.id = can.city\_id

where

c.title = 'Санкт-Петербург' and

can.age between 16 and 21

order by 2

*Выборка содержит список желаемых должностей, в основном в области программирования, разработки и системного администрирования. Питерская молодежь ориентирована на IT сектор, что вполне характерно для столичного или крупного развитого города.*

Общие выводы.

На основе анализа имеющегося датасета можно сделать некоторые выводы относительно характеристик рынка труда в РФ (исходя из предположения, что hh.ru объективно освещает весь рынок труда):

1. Основная часть активных соискателей находятся в возрасте от 20 до 45 лет. Как было уже отмечено ранее, в постсоветском пространстве исторически сложилась скрытая дискриминация по возрасту, порог 45-50 лет является “де факто” красной зоной, и сотрудников неохотно нанимают в этом возрасте. Хотя реалии уже давно изменились и специалисты хорошего качества в поиске нового места - это нормально, все же народ побаивается лишний раз пересмотреть свою позицию и условия за этой чертой. Отсюда и малое количество резюме в доступе для этой категории специалистов.

Доступ к рынок труда через агрегаторы подобные hh.ru ограничен для пожилых людей, в силу технологических сложностей использования. В перспективе 10-15 лет картина может резко измениться, так как на пенсию станут выходить граждане, для которых бурный рост информационных технологий пришелся на активную фазу жизни и карьеры, и для них использование этих каналов доступа не составляет труда.

1. По традиции для нашей страны, наблюдается высокая урбанизация рынков вообще и рынка труда в частности. Львиная доля активных соискателей (более 37%) сосредоточена в Москве, и оставшаяся большая часть в городах миллионниках, крупных отраслевых и производственных центрах, а также узлов сосредоточения мелкого и среднего бизнеса. В силу этой особенности, в столице и крупных городах создается наибольшее количество рабочих мест, что также усиливает приток рабочих и интеллектуальных ресурсов из провинциальных регионов.

Если более детально рассмотреть специфику видов должностей и работ, которые актуальны на этом рынке, то напрашиваются следующие выводы:

1. проектная работа популярна в Москве среди мужчин различных возрастов и следующих основных специальностей: программисты и разработчики, инженерных специальностей, менеджеры проектов или SMM-менеджеры, а также специалисты по системам бухучета типа 1С SAP и подобным. Существуют характерные типы специальностей, которые предполагают именно занятость в конкретных проектах на конкретный временной интервал, без привязки к штатному расписанию. Это также дает возможность трудоустроиться в Москве сначала на проект, а затем по возможности “закрепиться” и окончательно перебраться в столицу.
2. Среди молодежи в частности очень популярны специальности из сектора hi-tech и IT, а именно: разработчик, аналитик, программист. Наиболее желанные должности связаны именно с этими направлениями.
3. Выяснилось, что найти сотрудников для вахтового метода работы, например для предприятий нефтегазовой отрасли, достаточно проблематично. Связано это в первую очередь с тем, я полагаю, что в настоящее время культивируется тренд “работы из дома”. В частности, молодежь связывает зарабатывание денег с пребыванием в супер комфортных условиях за компьютером. В то время как, вахтовые работы, как правило связаны с долгосрочными командировками в места с суровыми условиями, связанными с отрывом от семьи, друзей и привычных благ цивилизации.