

Добрый день! Готова обратная связь по вашей работе.

Напомню критерии оценки проекта:

1. Правильность решения задач
2. Читабельность и верное форматирование кода, наличие комментариев
3. Ясность, доступность хода решения задач
4. Логичность и полнота выводов

По каждому из критериев работа оценивается максимум в 2 балла. Всего можно получить 8 баллов за проект.

Оценка вашей работы по вышеуказанным критериям:

1.Правильность решения задач - 2 балла.

Отличная работа! Всё верно.

Небольшие комментарии/рекомендации к запросам:

- Рекомендую сортировку по позициям полей в списке SELECT использовать только при проверке кода, а уже на этапе передачи кода на код-ревью или при фиксации кода в системе контроля версий писать алиасы полей. Часто в продакшене код с сортировкой по позициям отклоняют на код-ревью, так как это потенциальное место ошибки - изменится порядок полей и сортировка будет идти не по тому полю.
- Обратите внимание на поле `employment_type` таблицы `candidate`. Оно заполнено стандартно - все символы в нижнем регистре, значения разделены запятой с пробелом, поэтому применять `lower` к этому полю не стоит.
- Задание 3.3 также можно решить через `ILIKE` (поиск без учета регистра)
<https://postgrespro.ru/docs/postgresql/9.6/functions-matching#functions-like>
`desirable_occupation ilike '%разработчик%'`
`desirable_occupation ~* '%разработчик%'`
- В задании 3.5 условие `age < 100` повторяется, его лучше вынести за скобки и сделать общим, чтобы избежать дублирования кода.
- Обратите внимание на таблицу `timetable_type`. Она представляет из себя справочник с двумя полями - идентификатор типа и его наименование, которое содержит единственное значение в каждой строке. То есть использовать `like` в задании 4.1 для фильтрации вахтового метода не имеет смысла, здесь лучше обойтись простым оператором сравнения на равенство.

2.Форматирование кода и комментарии - 2 балла.

Все запросы хорошо отформатированы и читаемы. Во всех заданиях есть комментарии, но часто они излишние. Например, комментарии вида `-- таблицы городов и кандидатов связаны по ключу id города` можно смело исключать. Данные комментарии по сути просто дублируют код и не несут никакой дополнительной информации. В проекте большинство запросов простые и не требуют подробного комментирования, но наличие поясняющих

комментариев является одним из критериев оценки. Здесь было достаточно прокомментировать 3-5 спорных, сложных или интересных моментов в запросах.

3.Ход решения задач - 2 балла.

Все запросы построены понятно и логично. Весь изученный инструментарий использован обоснованно и уместно. Хочу отметить значимость критериев 2-3.

Всегда стоит помнить, что "код чаще читается, чем пишется", поэтому читабельность важнее "писабельности". Важно уметь писать понятный и удобочитаемый код.

4.Выводы - 2 балла.

К каждому заданию есть краткие, логичные выводы. Выводы корректно отражают существующие закономерности в данных. Сделан общий вывод по проекту. Корректно подмечены главные показатели набора данных. Выдвинуты логичные предположения по направлению дальнейшего анализа. Небольшое уточнение по столетнему кандидату: мы его не исключили из анализа в принципе, а только в тех случаях, когда идёт непосредственный анализ этого поля. На итоговом вебинаре по проекту я расскажу подробнее про обработку выбросов, присоединяйтесь.

По всем критериям получен максимальный балл. Замечательно!

Желаю вам дальнейших успехов в изучении Data Science.

Эталонное решение по проекту -

https://github.com/MarinaArh/sf_tutorials/blob/main/project_2_solution.ipynb

Итоговый созвон-вебинар по проекту состоится 3 сентября. Ссылка для подключения -

<https://us02web.zoom.us/j/85444465690?pwd=YzJlemhaUXBZV3NmSUIRU2RIS2xDQT09> Время 14-00, дополнительно будет анонс в канале проекта.

Если у вас остались вопросы по проекту, по оценке или любые вопросы по SQL - вы всегда можете мне их задать в чате проекта в Slack - #02-project-2.

С уважением, ментор проекта Марина Архипцева.