

## Модуль 4. Авиарейсы без потерь. Фидбек

Здравствуйте, Дмитрий!

Результаты вашей работы проверены. Сразу скажу, что работа заслуживает максимальное количество баллов по каждому пункту задания! Проверка подобных работ доставляет удовольствие. Все части выполнены последовательно и в полном объёме. Очень трудно найти ошибки и указать замечания к работе. Всё-таки я постараюсь дать несколько полезных рекомендаций в этом отзыве.

Решения задач из 4 и 5 раздела можно назвать близкими к идеальным. Большинство решений очень близки к авторским решениям, которые принято называть эталонными. Общее замечание, что не нужно переносить данные на новую строку. Не совсем правильно оставлять на строке только, например, SELECT. Форматирование кода и использование операторов можно считать оптимальным. Есть несколько рекомендаций. Например, задание 4.1, для которого в выводе можно было использовать оператор HAVING, позволяющий уточнить результат группировки:

```
SELECT a.city,  
       count(a.city)  
FROM dst_project.airports a  
GROUP BY a.city  
HAVING count(a.city) > 1  
ORDER BY 2 DESC
```

В задании 4.3.2 для объединения подзапросов можно было использовать оператор UNION (ALL):

```
SELECT 'Boeing' aircraft,  
       count(*)  
FROM dst_project.aircrafts a  
WHERE a.model LIKE 'Boeing%'  
UNION ALL  
...
```

Итоговый запрос. Подготовлен очень грамотный и содержательный запрос. Все нужные данные из условия задачи представлены в полном объёме. Есть только вопрос относительно столбца range. Будет ли полезна данная информация в последующем анализе? Вопрос дискуссионный. Все остальные данные полезны для последующего анализа. Из замечаний можно указать отсутствие комментариев, которые упрощают понимание отдельных частей запроса.

Дополнительное задание выполнено в блокноте jп. Можно сказать, что данная часть работы подготовлена лучше всего. Смею предположить, что было потрачено немало времени. Выполнен очень подробный анализ данных. Особенно хочется отметить использование корреляционного анализа. Многие

игнорируют возможность такого подхода. Хотя я считаю, что в данном случае от корреляционного анализа пользы мало. Вот если бы перед нами стояла задача поиска признаков для алгоритмов машинного обучения, то, да, он был бы полезен. Теперь касательно результатов. В целом, с ними можно согласиться. Фактически в нашем случае существенное влияние на итоговую прибыль оказывает стоимость потраченного топлива и количество проданных билетов на рейсе. Скажем так, что это предварительная оценка. Для полной оценки потребуются учёт иных расходов. Допустимо ограничиться только предварительной оценкой.

Презентация является логическим завершением отлично подготовленной работы. Все пункты из условия задачи представлены и раскрыты. Что меня поразило в презентации, так это аккуратность оформления. Все этапы работы приведены последовательно и имеют обязательные пояснения. Особо хочется отметить инфографику. Графики существенно упрощают понимание полученных результатов. В конце презентации указаны содержательные выводы и рекомендации, с которыми можно согласиться.

Если возникнут вопросы по отзыву, вы можете обратиться ко мне в Slack [@Сергей Добдин] в канале, соответствующем модулю с заданием/проектом [#0\_project\_4-работа-с-базой-данных].

Подведём итоги. Работу можно считать выполненной! Все основные пункты из условия задачи представлены. Считаю, что подготовлена отличная работа, которая заслуживает высокой оценки. Особое внимание заслуживает анализ данных в блокноте jп. Дмитрий, продолжайте изучать направление data science. Не бойтесь экспериментировать и творчески подходить к решению задач. Искренне желаю удачи в дальнейшем обучении!

Удачи в дальнейшем обучении!

Отзыв подготовил ментор @ Сергей Добдин

Буду ждать вас на итоговый созвон-вебинар по текущему проекту в пятницу в 20:00. На нём мы разберём основные ошибки группы и посмотрим эталонное решение. До встречи на вебинаре!

Ссылка на вебинар –

<https://us02web.zoom.us/j/85341744801?pwd=RFlidzRaeWUrVFNLrZj0YUxCd2xsdz09>