

Добрый день!

Практически по всем пунктам критериев Вы правильно выполнили проект, отличная работа!

- + оформление в Jupyter Notebook в соответствии с ноутбуком-шаблоном;
- + структура оформления (отформатированные выводы в отдельных ячейках типа Markdown, хорошо оформленный лаконичный код, ячейки сделали очень наглядными, отлично, что выделили отдельно выводы - удобно и понятно разбирать Ваше решение);
- + широкое использование пройденных в курсе библиотек, ничего лишнего;
- + читаемый и понятный код, отдельно хочется отметить грамотно продуманные имена переменных и функций;
- + оформили графики по всем правилам, плюс за содержательные названия и подписи осей.

В разделах "Исследование структуры данных" и "Преобразование данных" все задания выполнены верно.

В разделе "Исследование зависимостей в данных", к сожалению, на github не отображается ни один график: (Это можно решить следующими способами:

- В аргументе `fig.show()` прописывать какое-нибудь имя (тогда картинка сохранится) или использовать функцию `write_image` (примеры можно посмотреть по ссылкам:
https://github.com/nicolaskruchten/plotly_img_test/blob/master/Untitled.ipynb,
https://github.com/nadiakova/DS-learning/blob/main/churn_nk.ipynb)
- Использовать сервис <https://nbviewer.org>

Можно также сделать поподробнее выводы - добавить кроме статистических численных данных больше закономерностей и предположения о том, почему могли быть получены такие результаты (особенно в последних заданиях). Например:

- во 1-м мода распределения равна примерно 30 годам. Такие результаты могут быть получены вследствие того, что в этом возрасте люди чаще всего имеют желание сменить работу или даже род деятельности после того как, например, поработали после окончания ВУЗа по специальности. В более старшем возрасте работники уже либо идут на повышение, либо получают офферы, т.е. у них нет необходимости пользоваться такими ресурсами как HH.
- Можно также отметить корреляцию первого и второго графиков.
- в 6-м готовность к переезду и командировкам сильно повышает уровень заработной платы. Вероятно это связано с тем, что командируемые - это, как правило, люди занимающие высокие должности;
- в 7-м для категорий "среднее специальное" и "среднее" образование медианная заработная плата слабо растет с увеличением возраста, в отличие от высшего образования и неоконченного высшего;
- в 8-м наблюдается прямая зависимость опыта работы от возраста. Точки лежащие на прямой и находящиеся выше нее - аномальные значения опыта работы, равного или превышающего возраст соискателя.

В разделе "Очистка данных" всё правильно.

Отличные дополнительные графики!

Отдельно хочется отметить:

Здорово, что написали функции во многих заданиях для выделения признаков и оформили комментарии.

Плюс за лаконичность кода и широкое применение методов и функций из пройденных в курсе библиотек.

Отдельный плюс за исправленную нумерацию заданий и за выделенные выводы к графикам.

Отличное оформление на github!

Подробный разбор всех заданий будет в формате видеозвонка, в субботу в 11:00, ссылка

<https://us02web.zoom.us/j/85882453347?pwd=b2E5cUdxTU1kWFNhbXptYXlmWWoyZz09>

(также будет анонс в чате Вашей группы).

Отзыв подготовила ментор Мария Жарова. Если возникнут вопросы, можете обратиться ко мне в канал # 01-project-1 в Slack, постараюсь на всё ответить и разобраться с моментами, которые вызывают трудности. Удачи в обучении!