Задачи

1. Импортируйте библиотеку pandas как pd. Загрузите два датасета user\_data и logs. Проверьте размер таблицы, типы переменных, наличие пропущенных значений, описательную статистику.
2. Какой клиент совершил больше всего успешных операций? (success == True)
3. С какой платформы осуществляется наибольшее количество успешных операций?
4. Какую платформу предпочитают премиумные клиенты?
5. Визуализируйте распределение возраста клиентов в зависимости от типа клиента (премиум или нет)
6. Постройте график распределения числа успешных операций
7. Визуализируйте число успешных операций, сделанных на платформе computer, в зависимости от возраста, используя sns.countplot (x – возраст, y – число успешных операций). Клиенты какого возраста совершили наибольшее количество успешных действий?

Описание данных

[user\_data](https://yadi.sk/d/gvlTf9t0ewyoaA):

* **client** – идентификатор пользователя
* **premium** – является ли клиент премиум
* **age** – возраст

[logs](https://yadi.sk/d/Dc6O9jHAGfWgyQ):

* **client** – идентификатор пользователя
* **success** – результат (успех - 1, нет - 0)
* **platform** – платформа
* **time** – время в формате Unix

Импортируйте библиотеку pandas как pd. Загрузите два датасета user\_data и logs. Проверьте размер таблицы, типы переменных, посмотрите на описательную статистику.

**Вопрос**: Сколько уникальных значений принимает переменная platform?

[Подсказки](https://www.notion.so/3-1-python-b8bd15dcff0e49b599eb7ea482ffb083)

Какой клиент совершил больше всего успешных операций?

Если несколько пользователей совершили одинаковое число успешных действий, перечислите их идентификаторы (client) через запятую и пробел в порядке возрастания. Например:

12, 123, 456

[Подсказки](https://www.notion.so/3-2-python-8a46ce60f9bb44d6a2cfa7dd27574f65)

С какой платформы осуществляется наибольшее количество успешных операций?

[Подсказки](https://www.notion.so/3-3-python-d0dd443a014c4df5ae7bae884f413bd1)

Phone

Tablet

Web

Computer

Какая платформа наиболее популярна среди премиумных клиентов?

[Подсказки](https://www.notion.so/3-4-python-5553be7e7c2b45cba18d2f3281c38f85)

Web

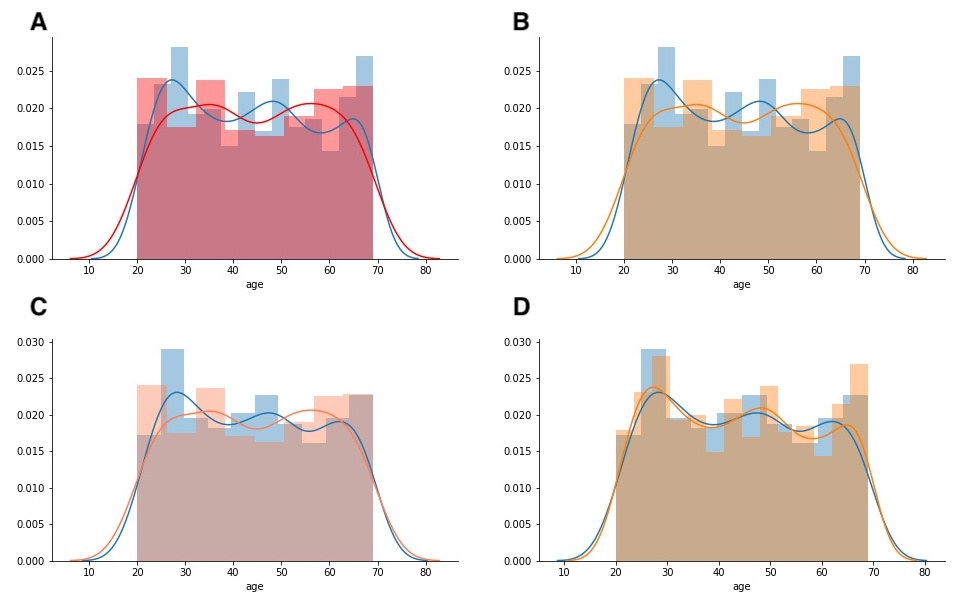
Phone

Tablet

Computer

Визуализируйте распределение возраста клиентов в зависимости от типа клиента (обычный или премиум), используя библиотеку seaborn с параметрами цвета **по** **умолчанию**. Какой график получится в результате?

Используйте датасет, который вы получили на предыдущем шаге



[Подсказки](https://www.notion.so/3-5-python-21c3110ea15042308180b76196d7f6c0)

A

B

C

D

Постройте график распределения числа успешных операций:

* *x* – число успешных операций
* *y* – число клиентов, совершивших столько успешных операций

[Подсказки](https://www.notion.so/3-6-python-03b5afb5ca944c2388817047830afee9)

Выберите подходящие утверждения:

Все клиенты совершили не более 5 успешных операций

Несколько клиентов совершили более 40 успешных действий

Бóльшая часть наблюдений сосредоточена в правой части графика

Бóльшая часть наблюдений сосредоточена в левой части графика

⭐️Задание со звёздочкой!⭐️

Визуализируйте число успешных операций, сделанных на платформе computer, в зависимости от возраста, используя sns.countplot. По оси *x* – возраст, *y* – число успешных операций. Если размер графика получается небольшим, можно добавить plt.figure(figsize=(12, 8)) в начало соответствующей ячейки.

Полезные ссылки:

* [sns.countplot](https://seaborn.pydata.org/generated/seaborn.countplot.html) – документация к типу графика
* [figsize](https://lab.karpov.courses/learning/212/module/2218/lesson/20491/59038/277294/) – про размер графиков

**Вопрос**: Клиенты какого возраста совершили наибольшее количество успешных действий?

[Подсказки](https://www.notion.so/3-7-python-1a3da845619945af8ba7046d899f35a3)