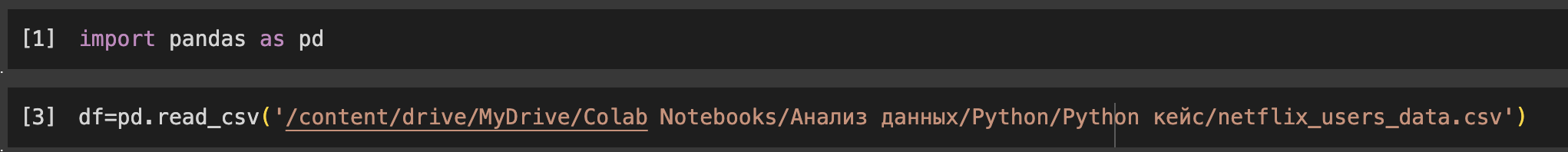
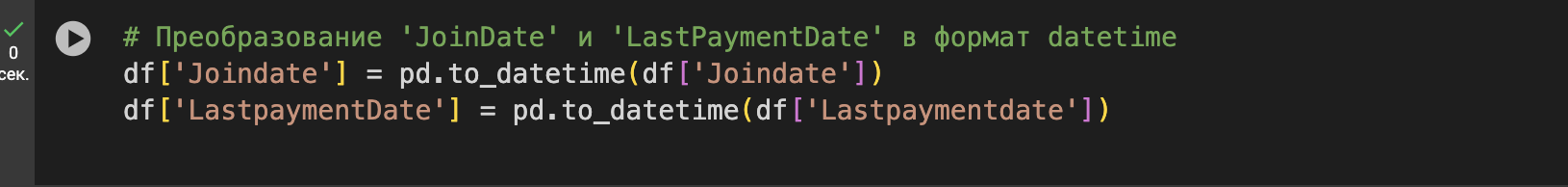
**Phyton Урок 14. Кейс Netflix**

Рассмотрим второй кейс — данные Netflix. В этом датафрейме есть колонки user\_id, тип подписки, месячная выручка, дата активации пользователя, последняя дата совершения платежа, страна, возраста пользователя, пол, тип девайса и продолжительность подписки. В этом кейсе базовые действия те же самые, что и в предыдущем задании.

При передаче данных в ChatGPT нужно сообщить, в каком формате содержатся исходные данные: тогда он задаст формат исходного текста. Например, формат даты может быть передан как через косую черту, так и через дефис, и т.д.

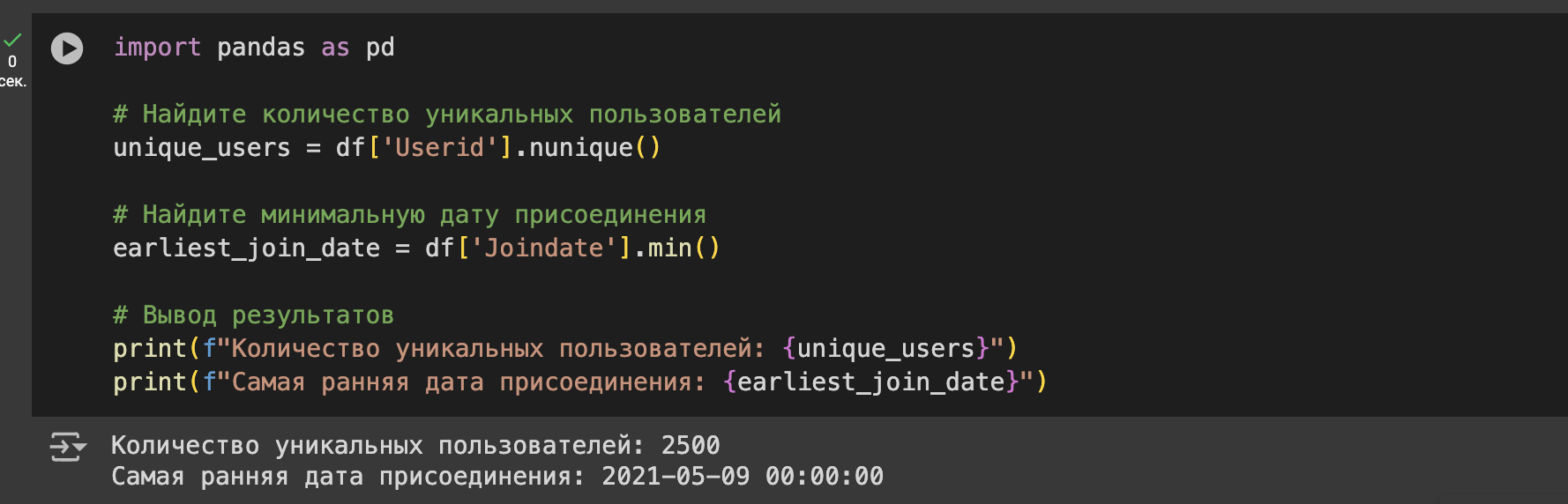




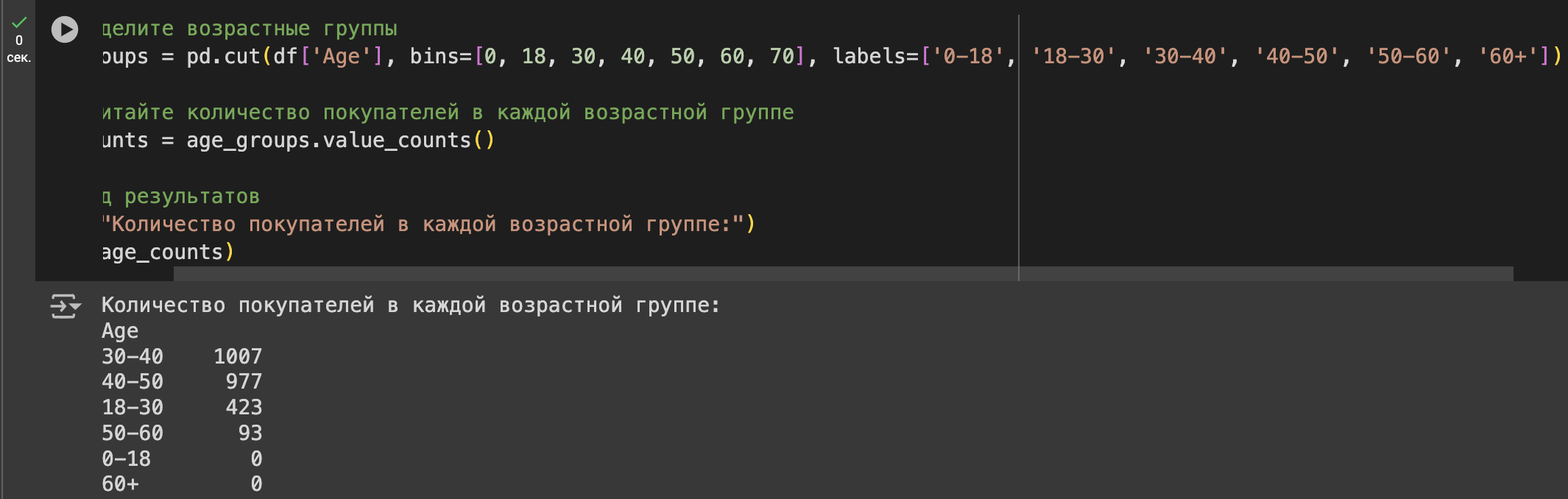


В основной части исследовательского анализа выполни следующие действия:

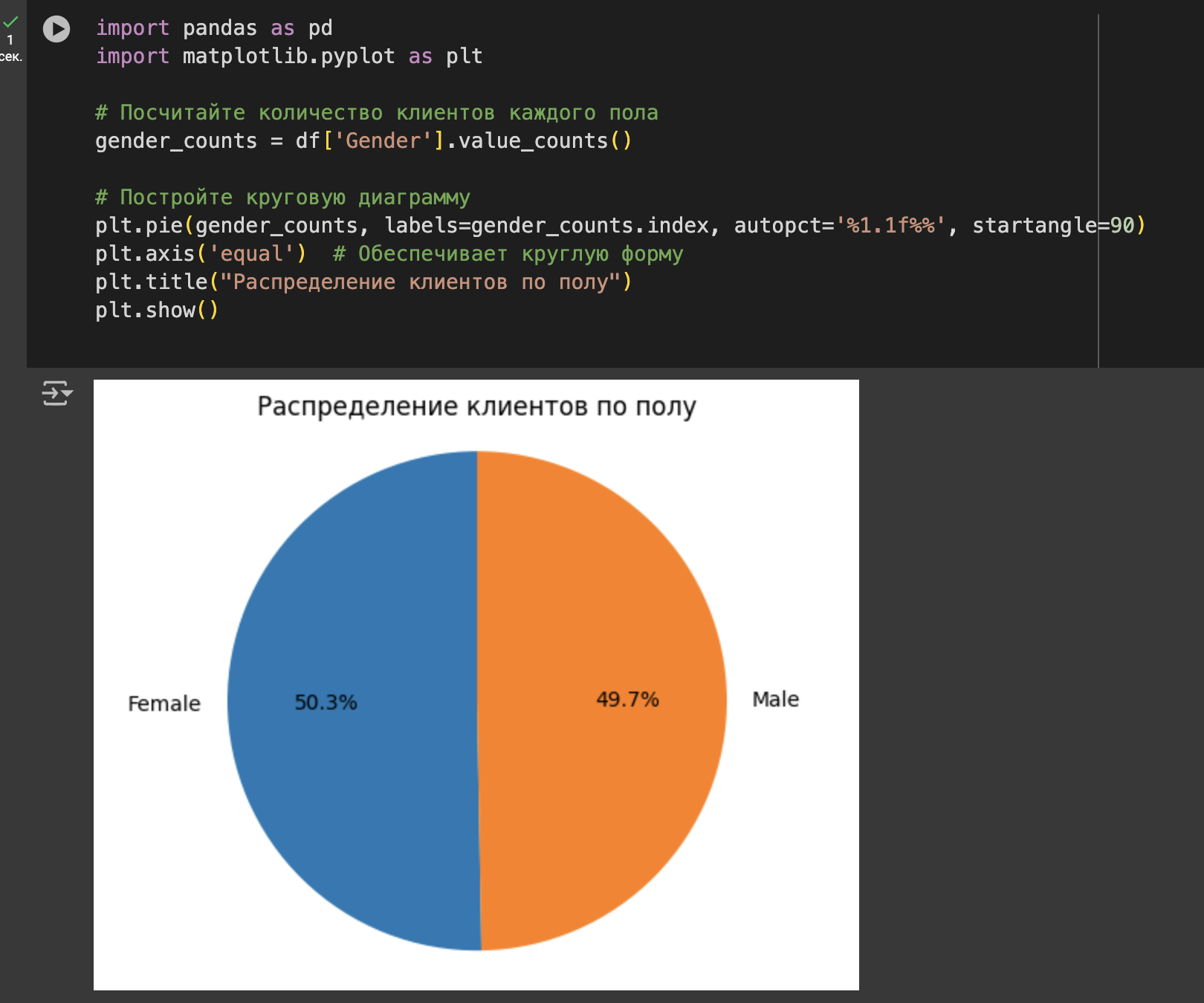
* узнай количество уникальных пользователей и период, с которого появились эти пользователи;



* изучи возраст пользователей;



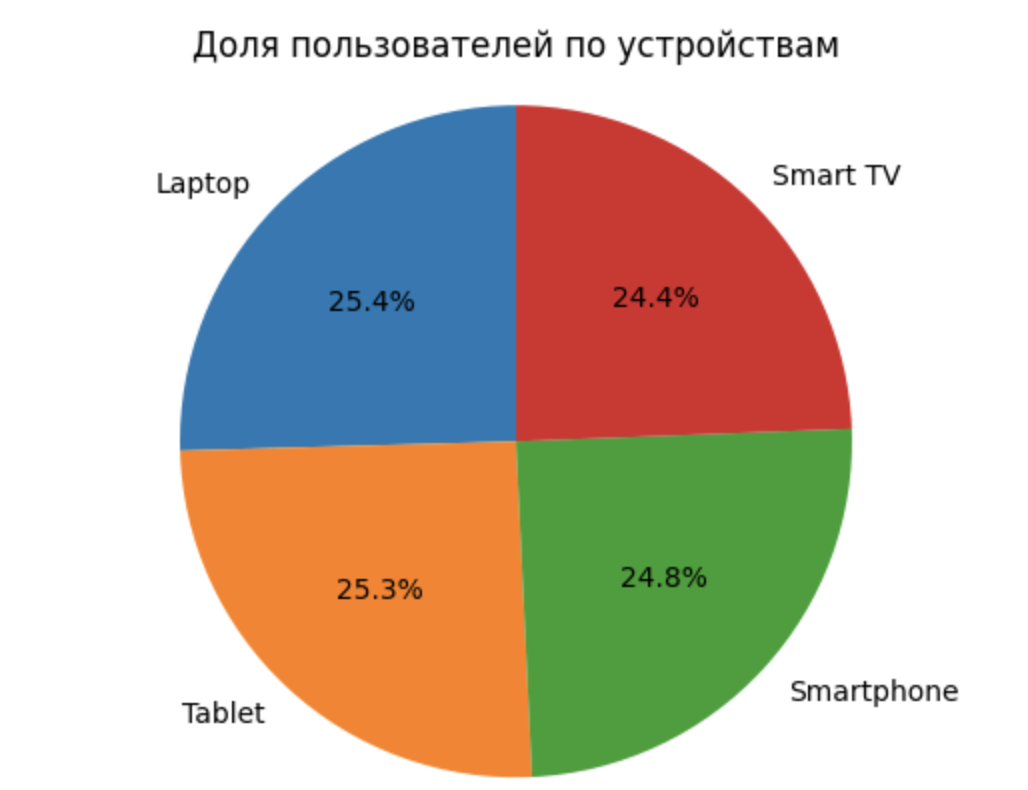
* изучи пол пользователей;

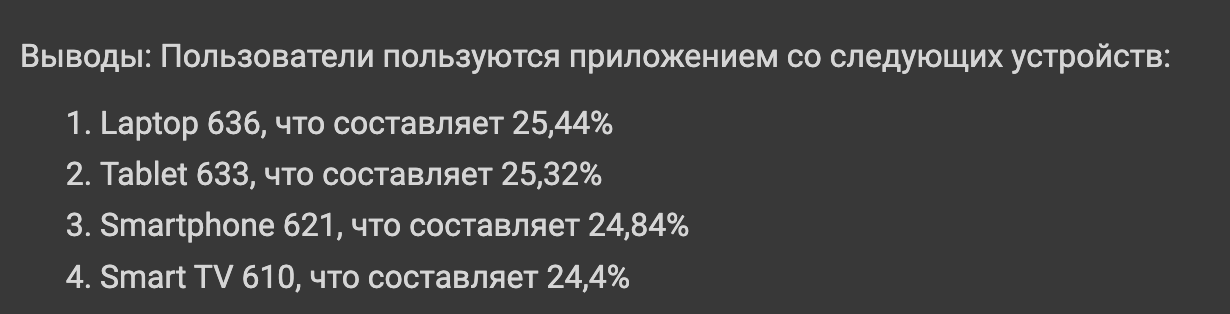


* изучи девайсы пользователей (можно визуализировать этот параметр через столбчатую либо круговую диаграмму);

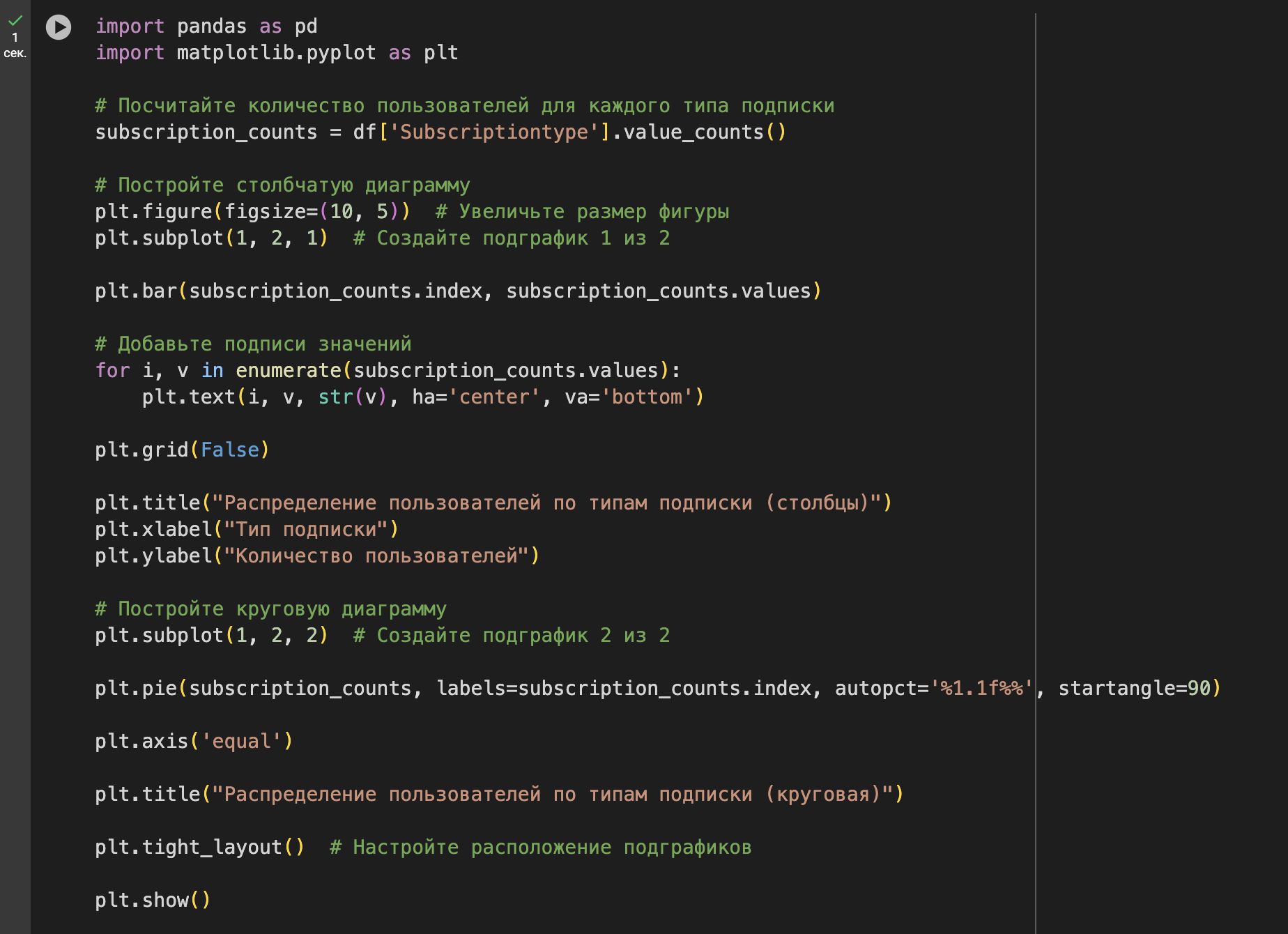


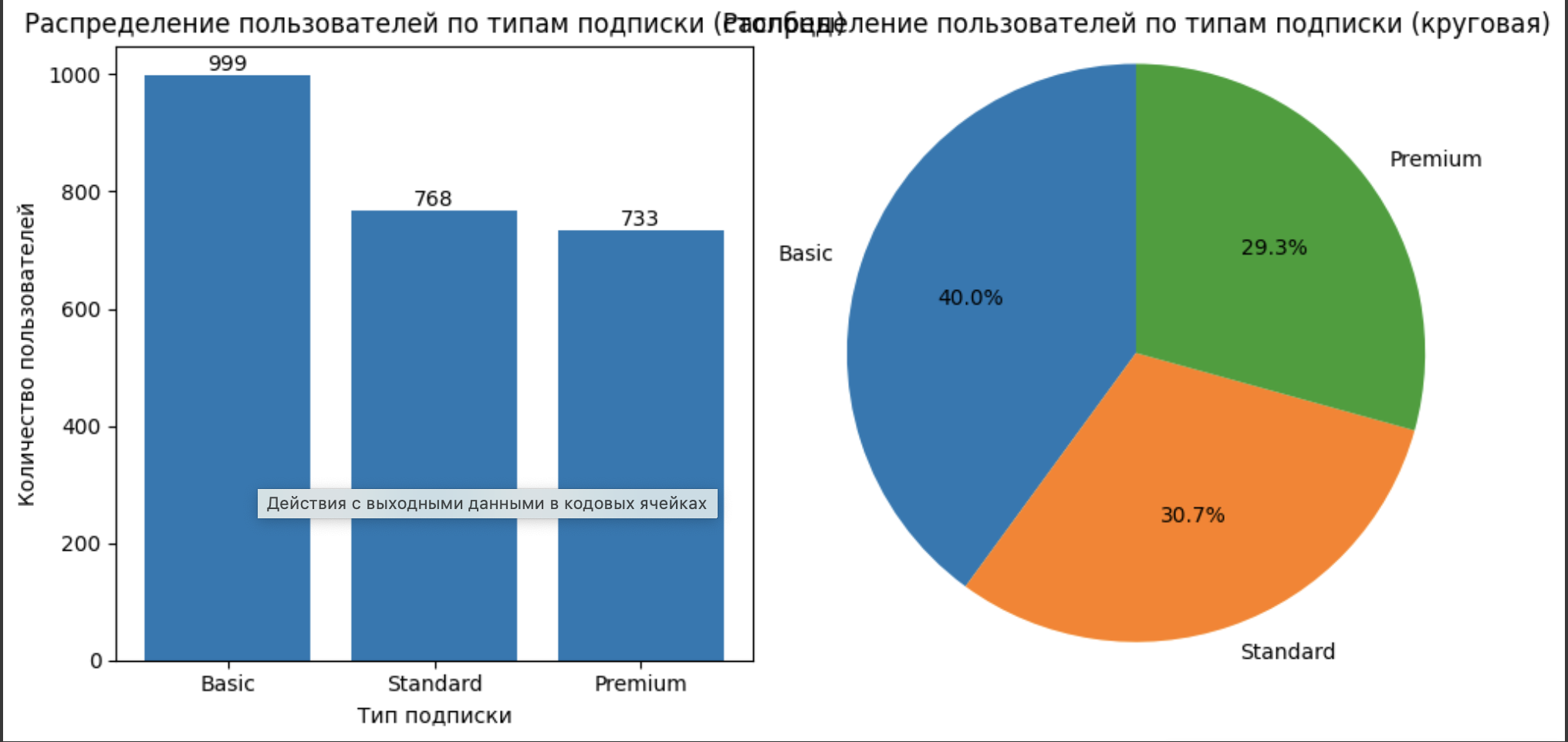


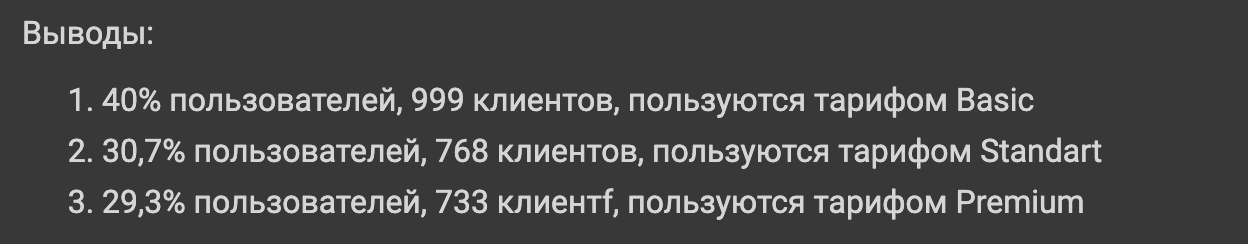




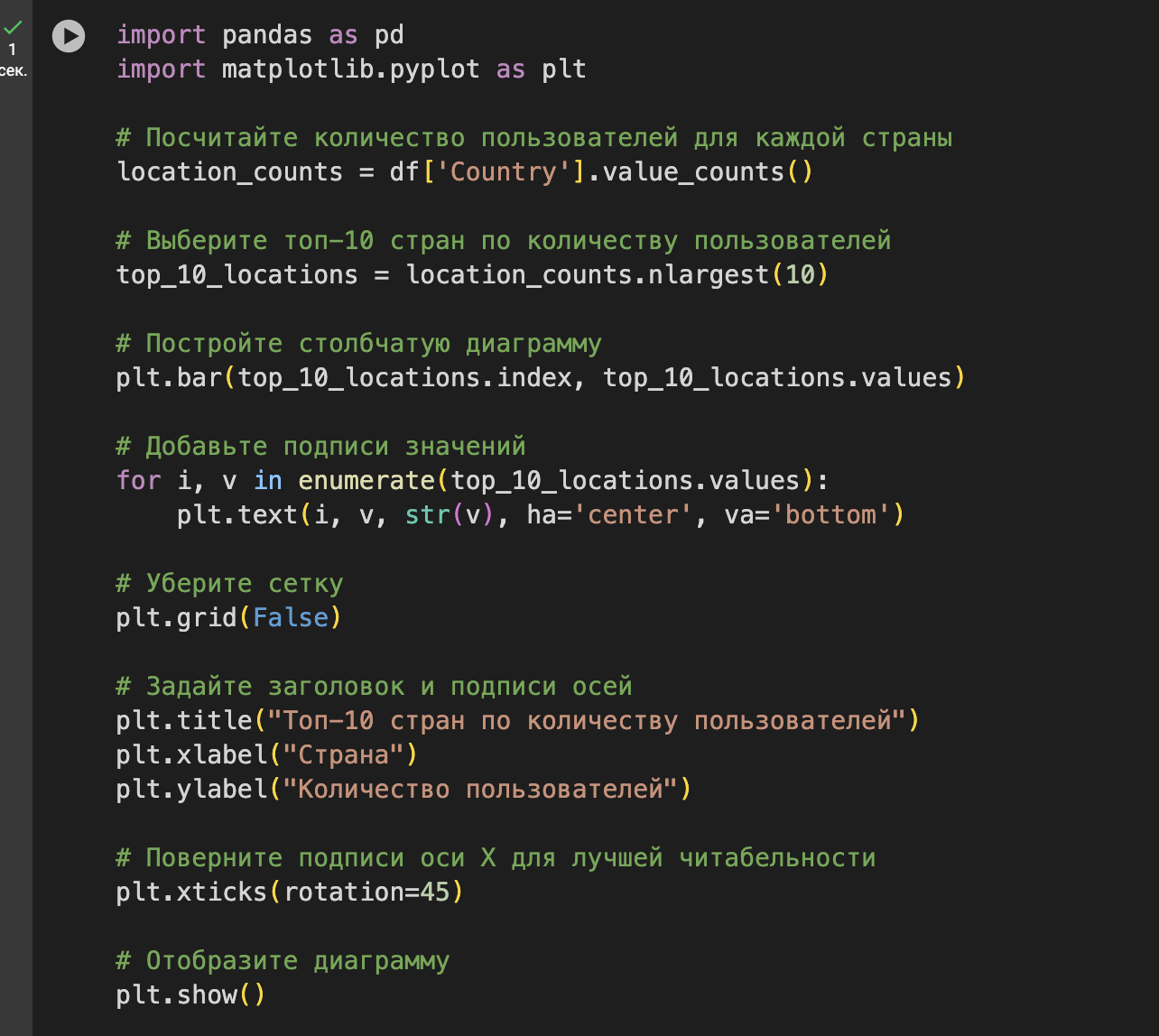
* изучи тип подписки пользователей. Не забудь написать выводы;

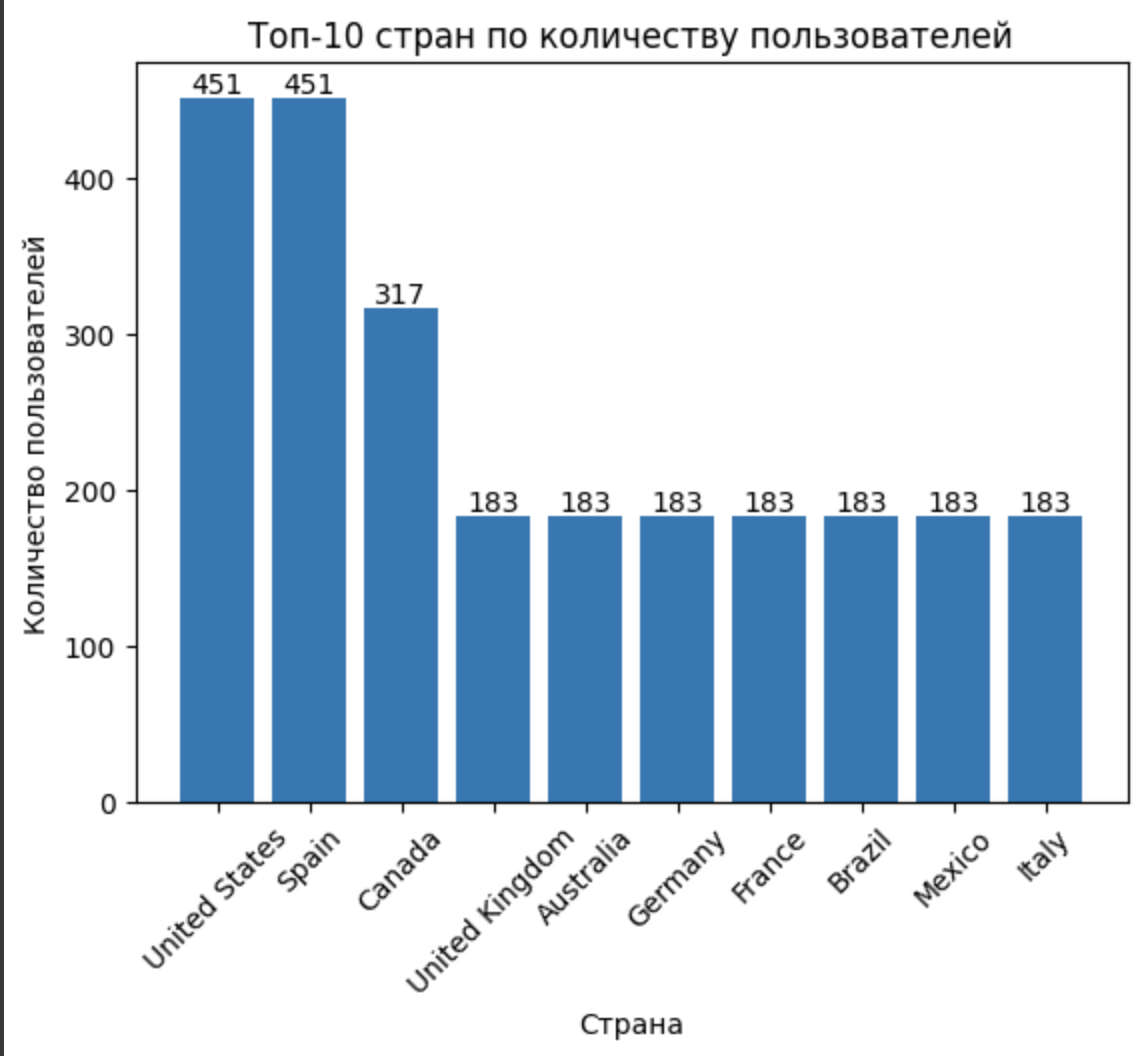


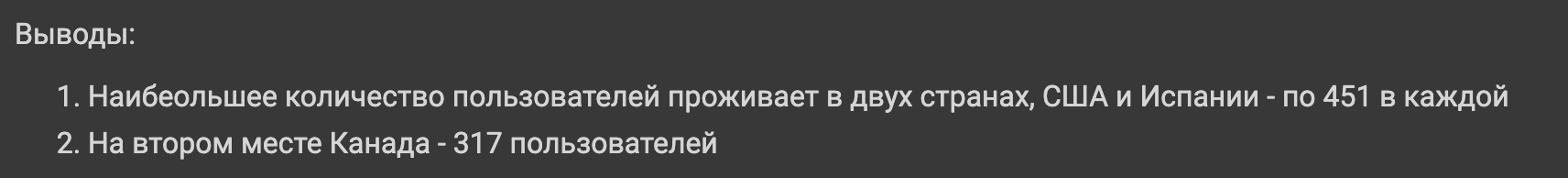




* изучи локацию пользователей на основе информации о стране проживания;



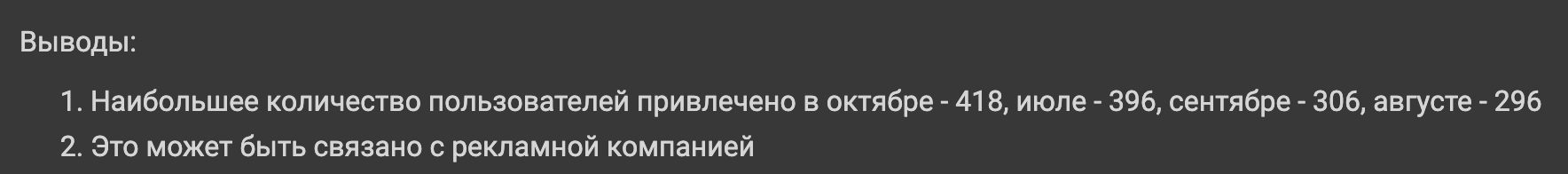




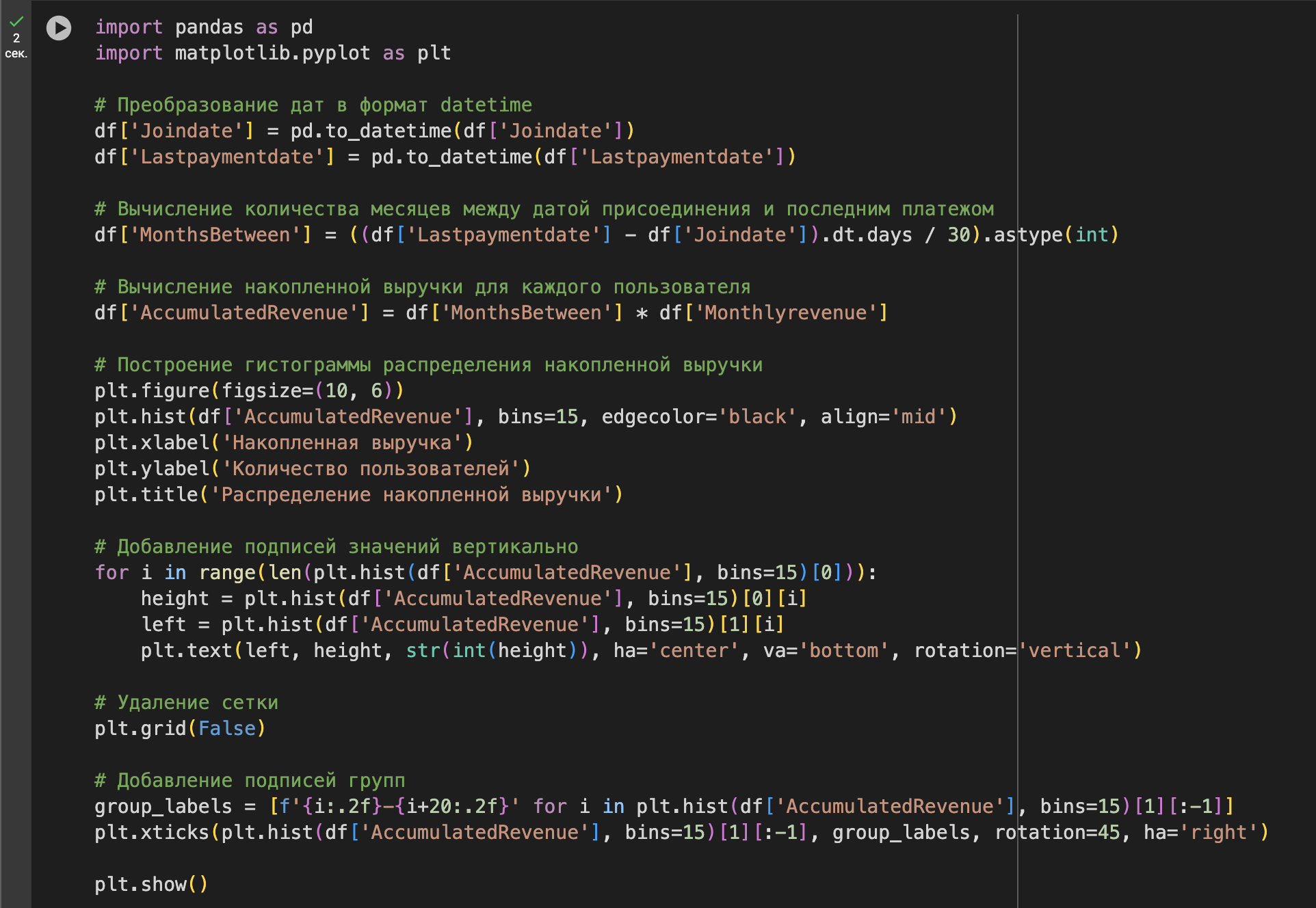
* построй линейный график динамики привлечения пользователей на основе данных Join date. Если график будет слишком мелким, сгруппируй данные по месяцу;

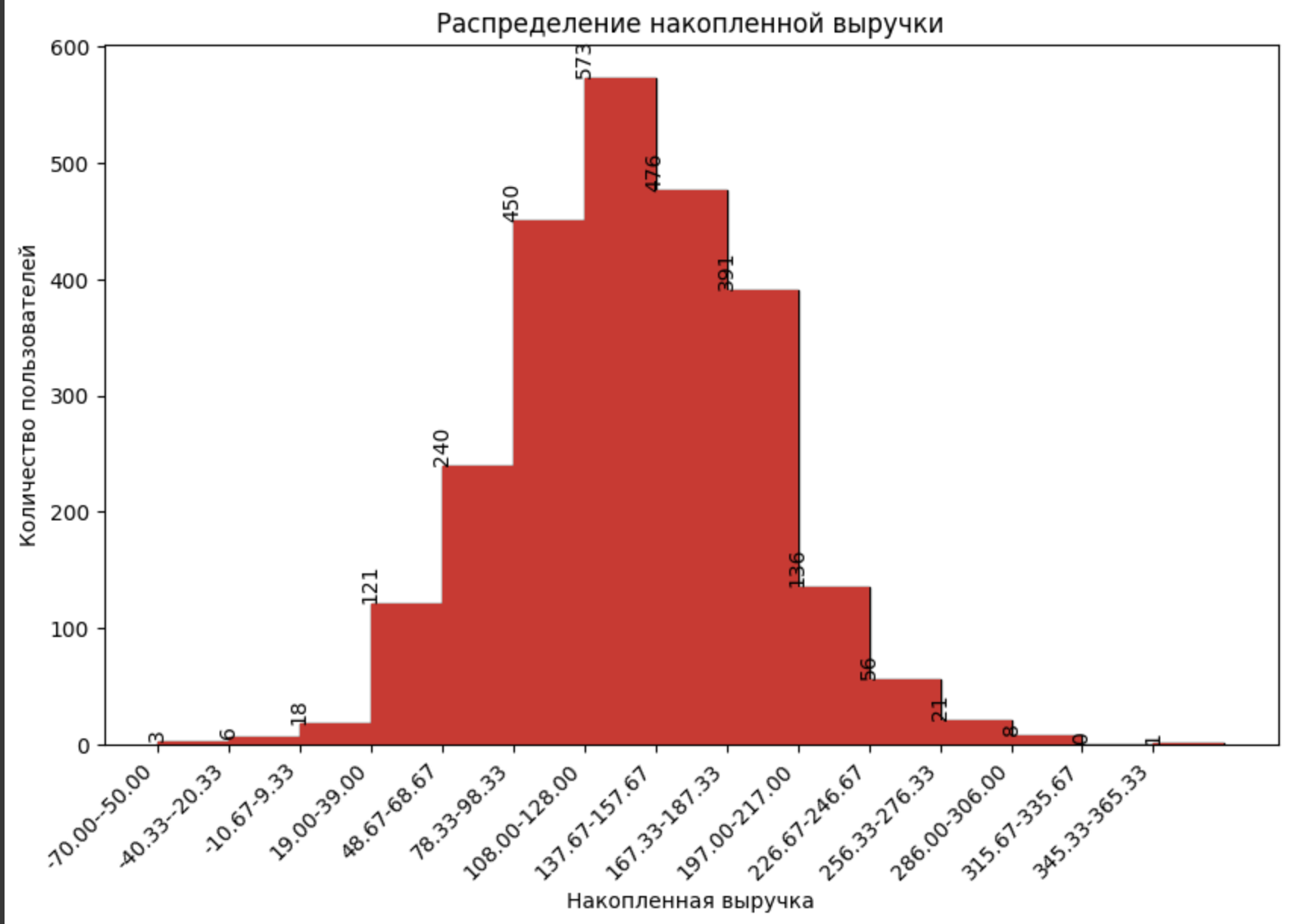


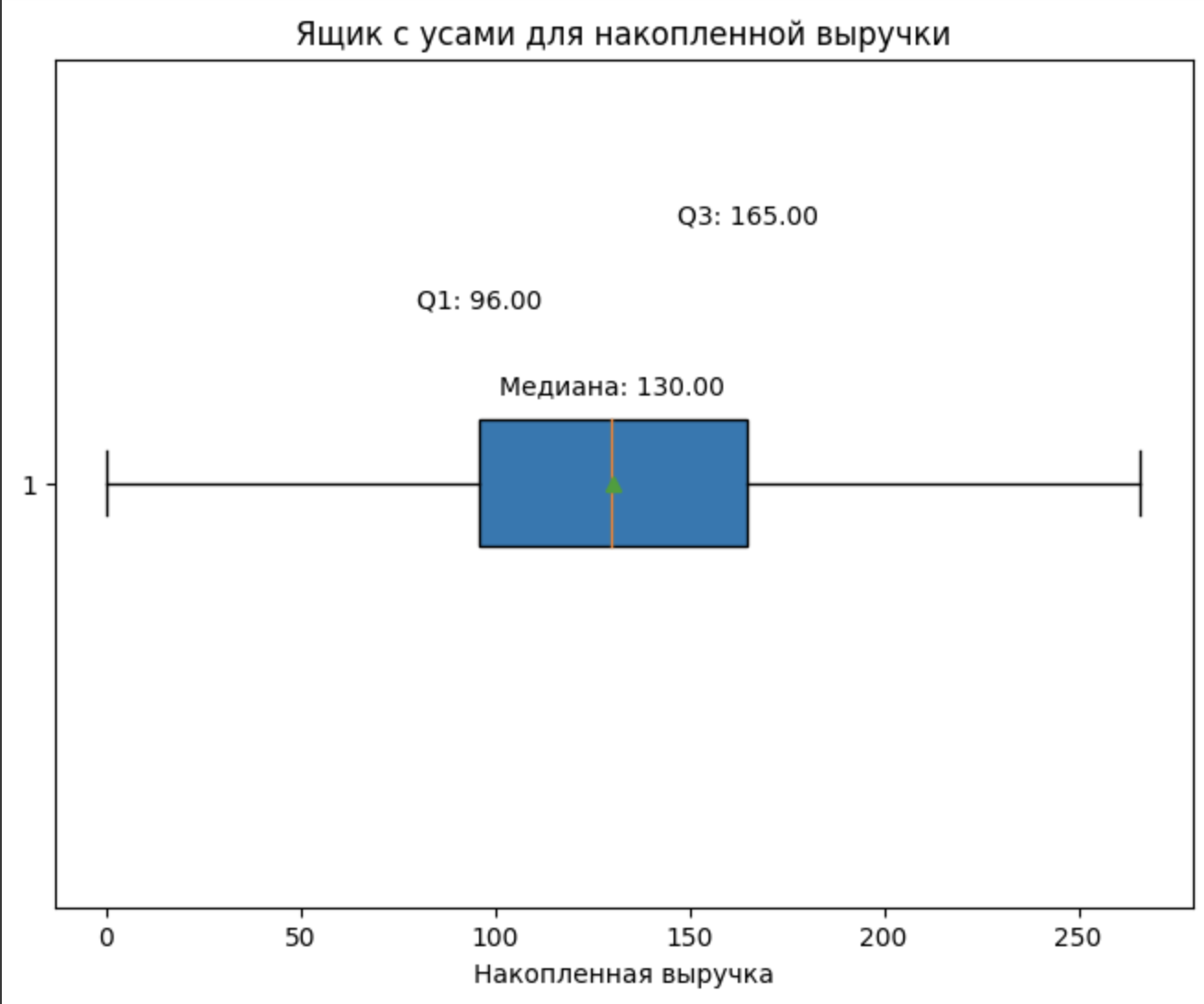


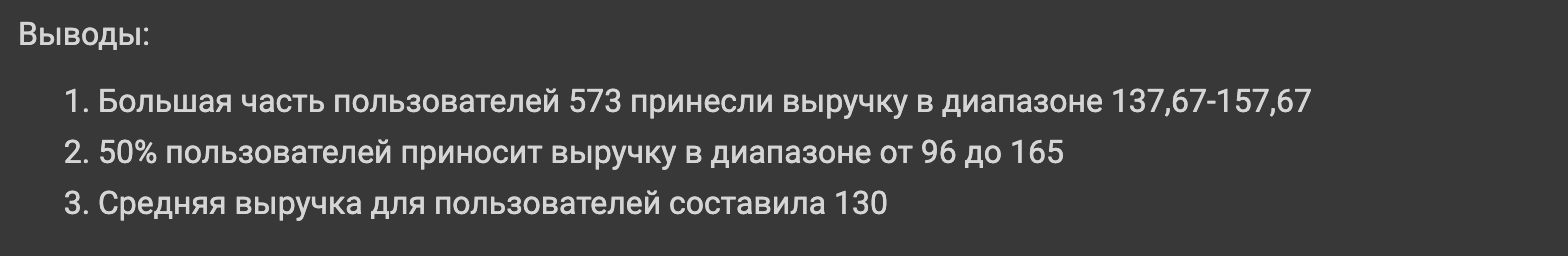


* изучи динамику выручки сервиса - необходимо взять накопленные значения Monthly Revenue в промежутке между Join date и Last payment date, поскольку в датафрейме есть пользователи, которые платили, но потом ушли из сервиса, и есть пользователи, пришедшие позже других. Нужно узнать, сколько прошло времени между первой датой платежа и датой старта, умножить на ежемесячный платеж и проанализировать накопленное значение по каждому пользователю, после чего посмотреть на общую выручку сервиса;

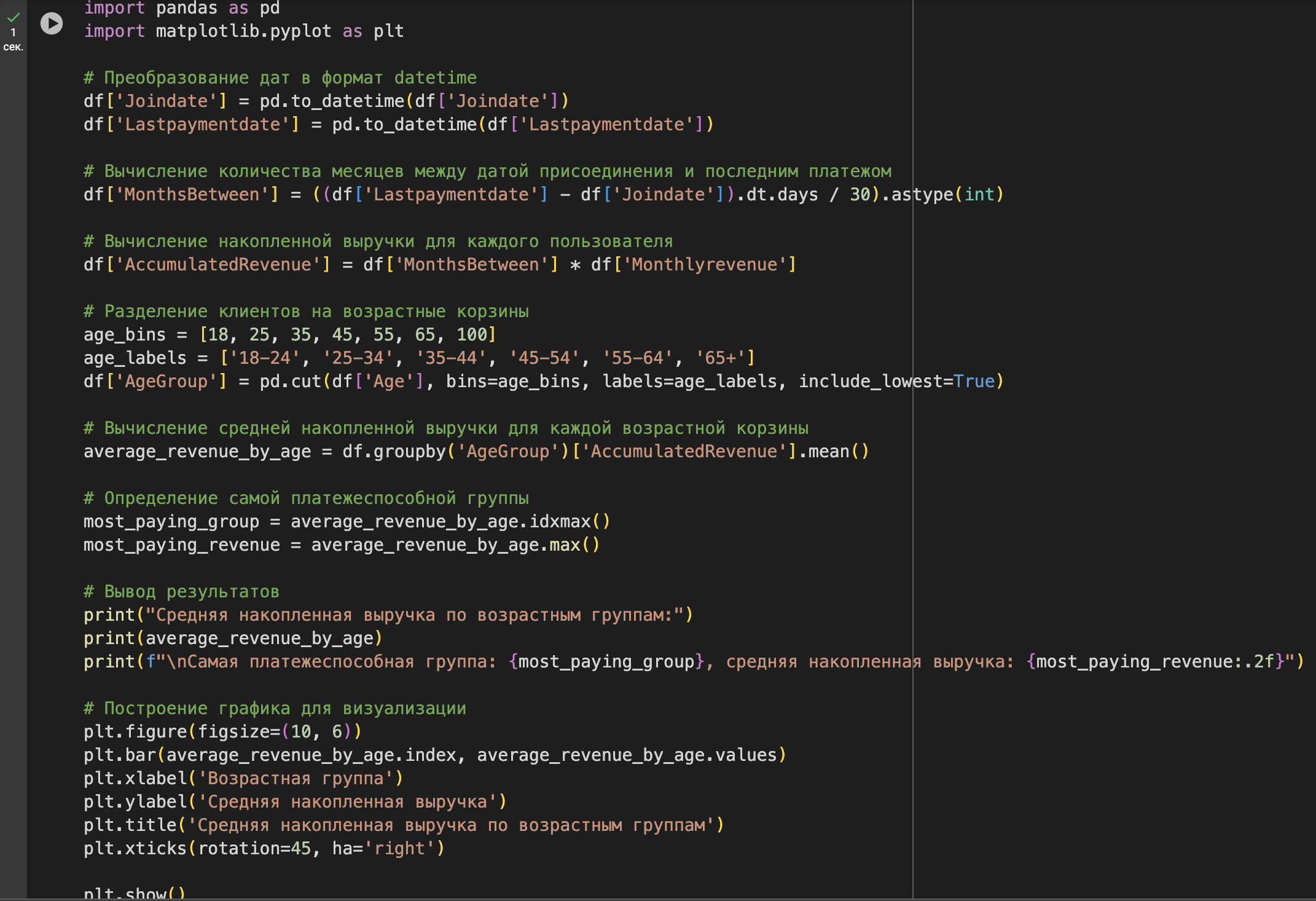


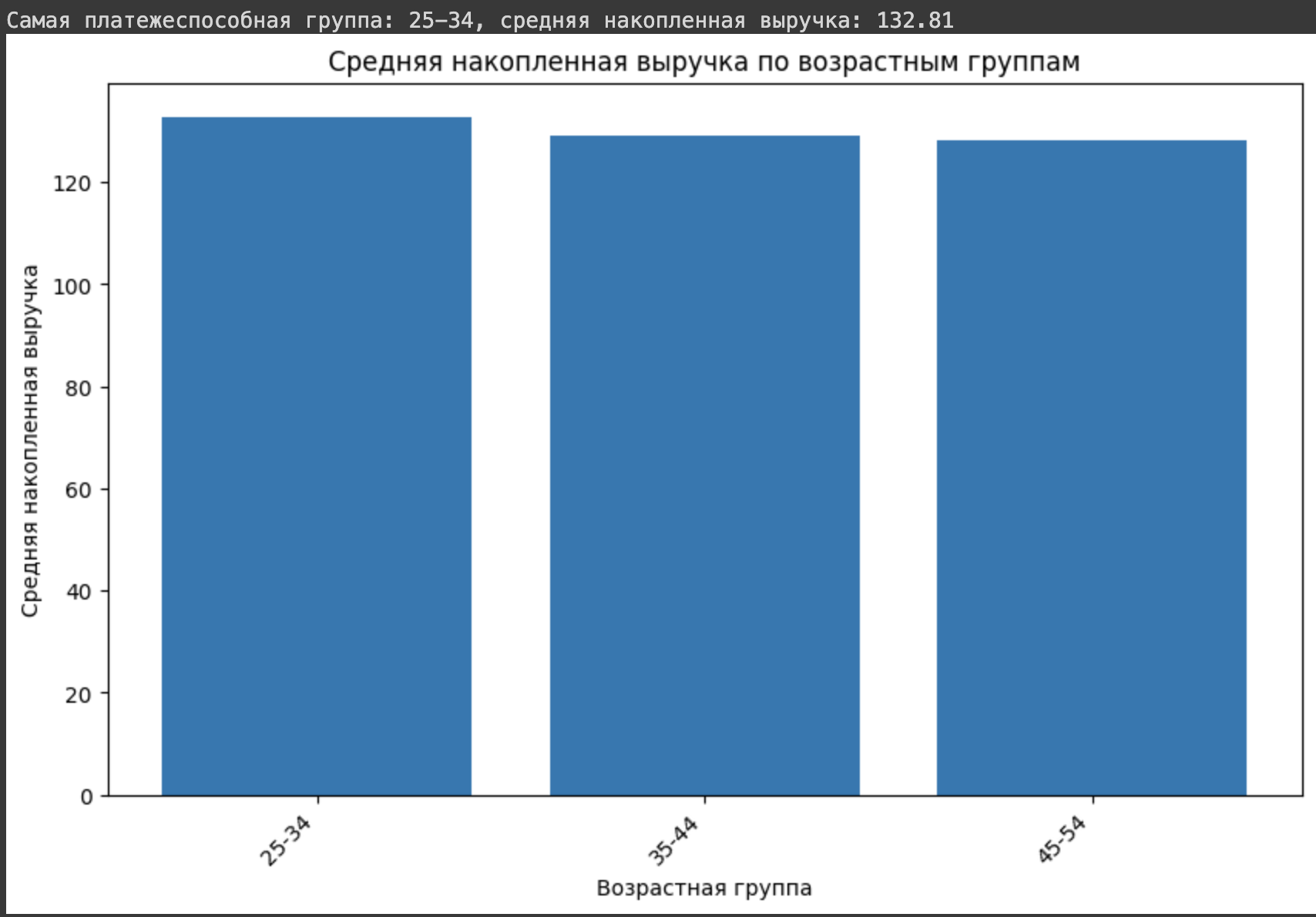


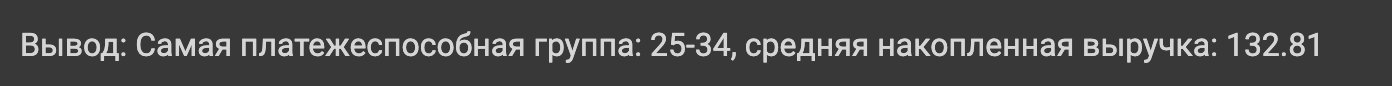




* раздели клиентов на корзины по возрасту и определи самую платежеспособную группу;

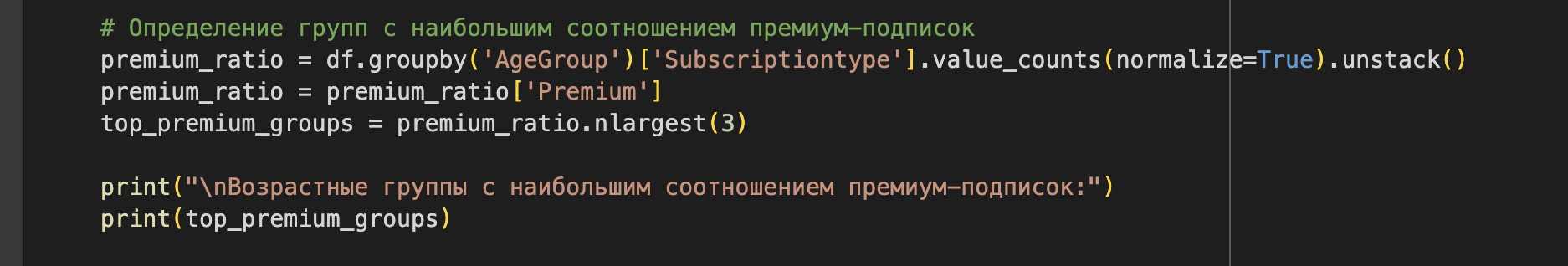


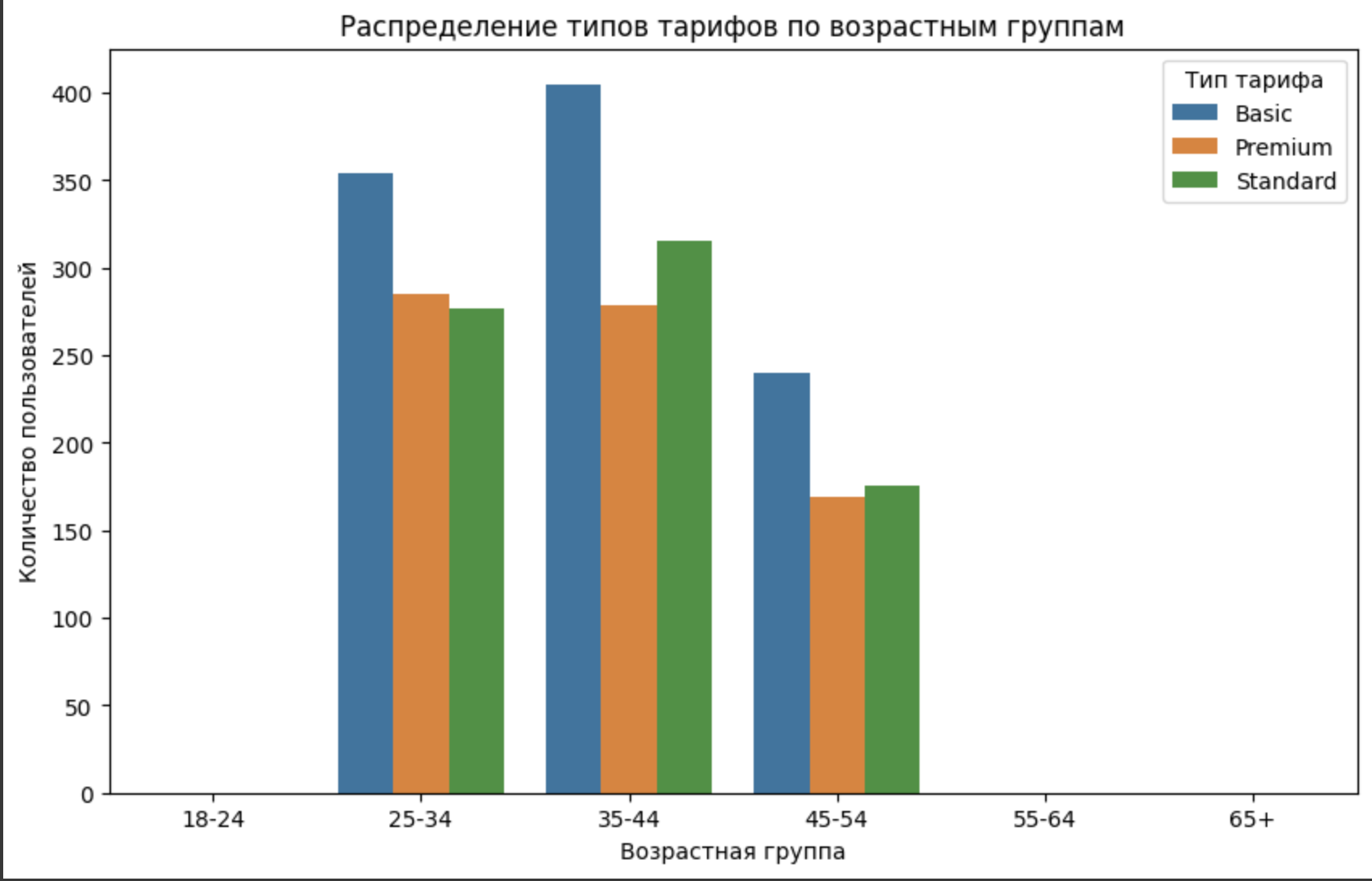


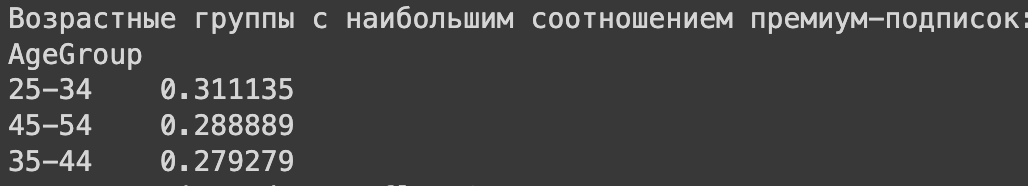


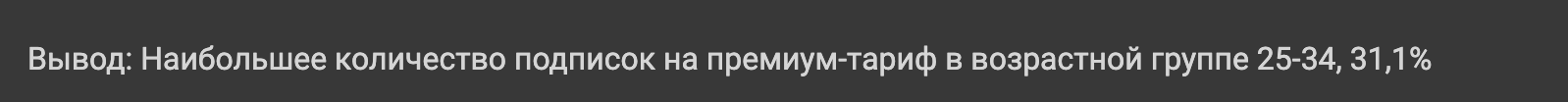
* выяви зависимость типа тарифа от группы возраста пользователя - возможно, в каких-то группах по возрасту есть преимущественные тарифы, и таким образом можно определить, какая аудитория более платежеспособна. При этом можно посмотреть корреляцию или построить графики, анализируя внутри каждой группы разбивку под Subscription Type. Далее выдели группы, где соотношение, допустим, премиум-подписки будет наибольшим.



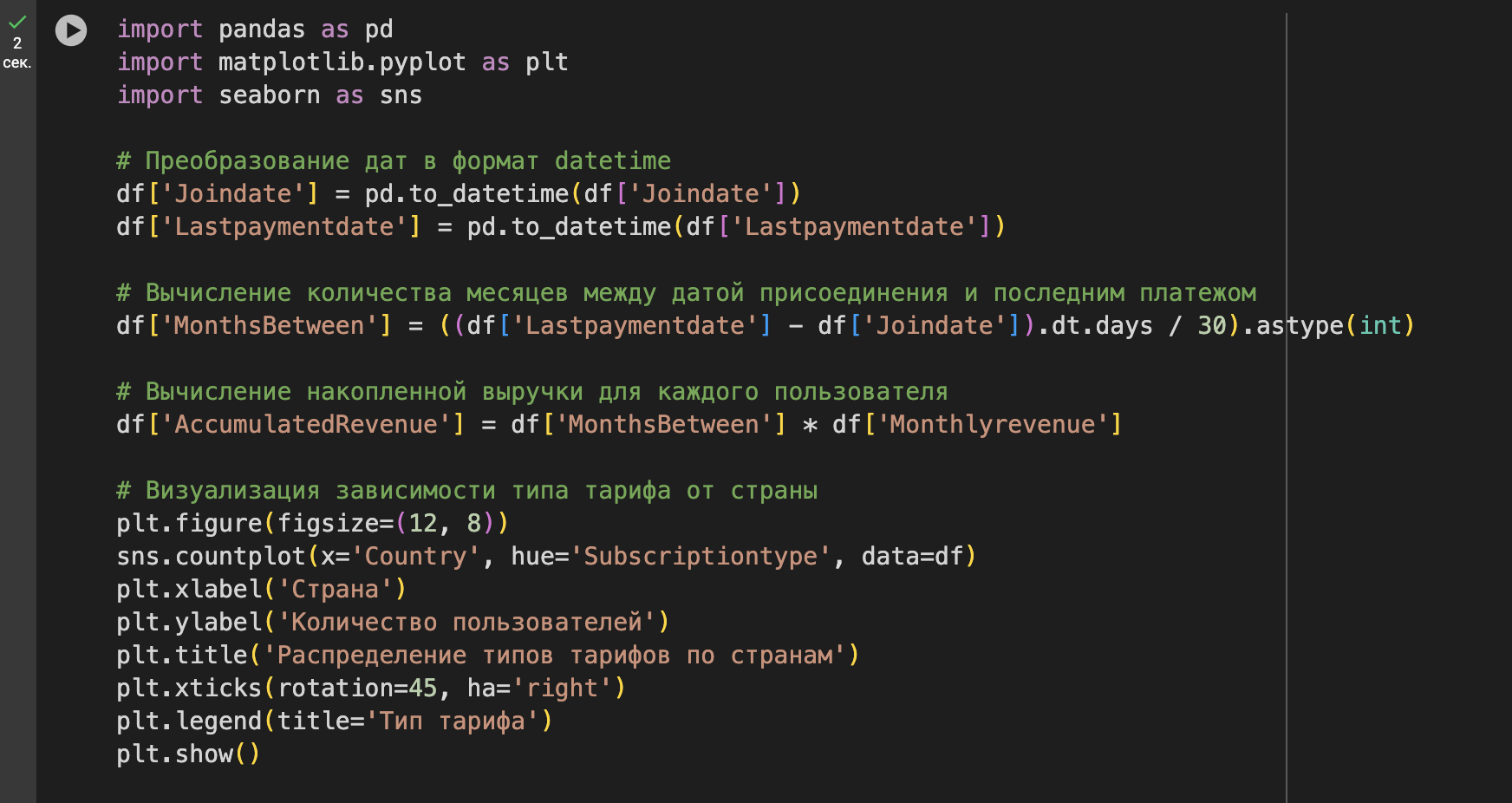


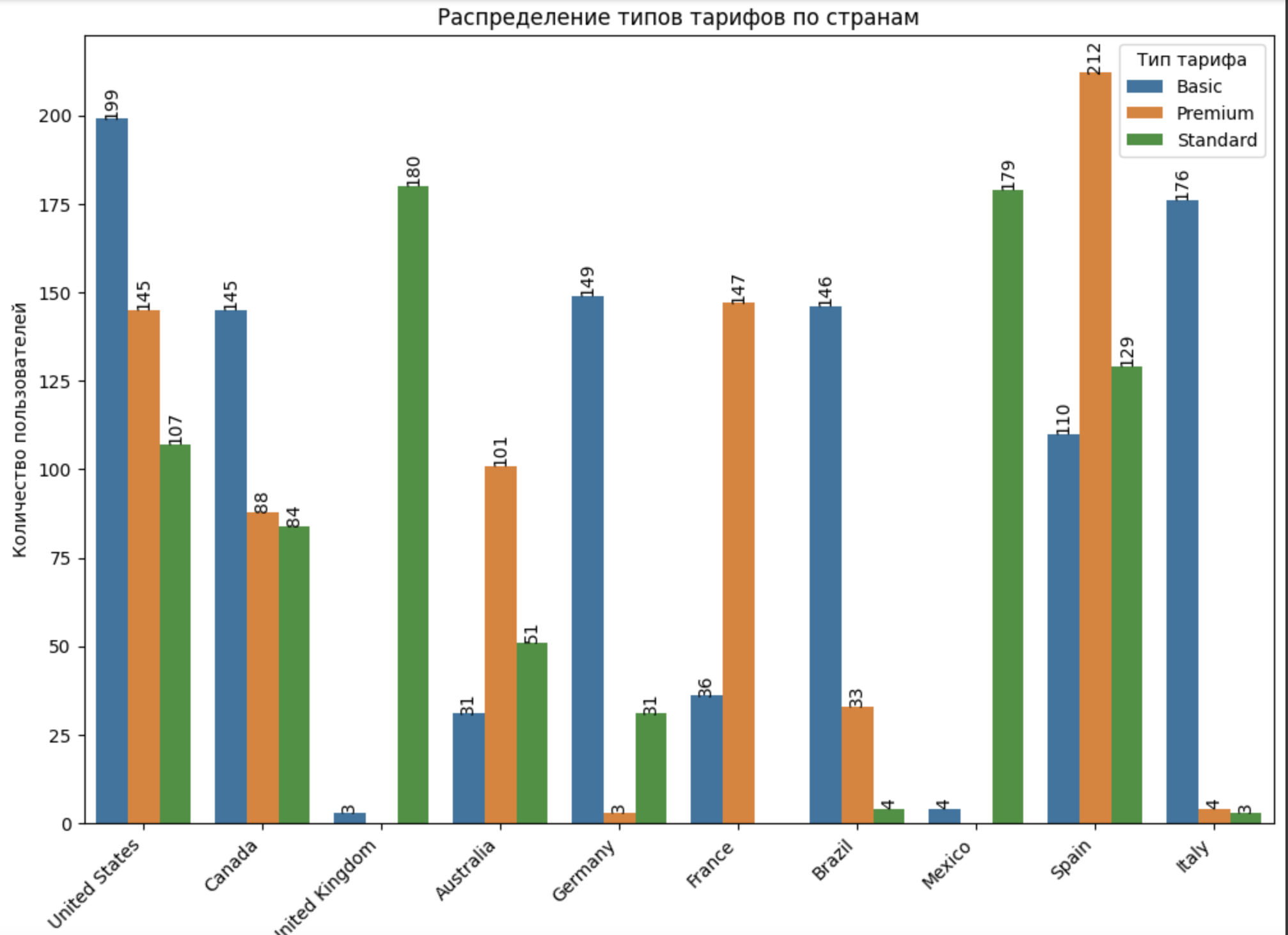


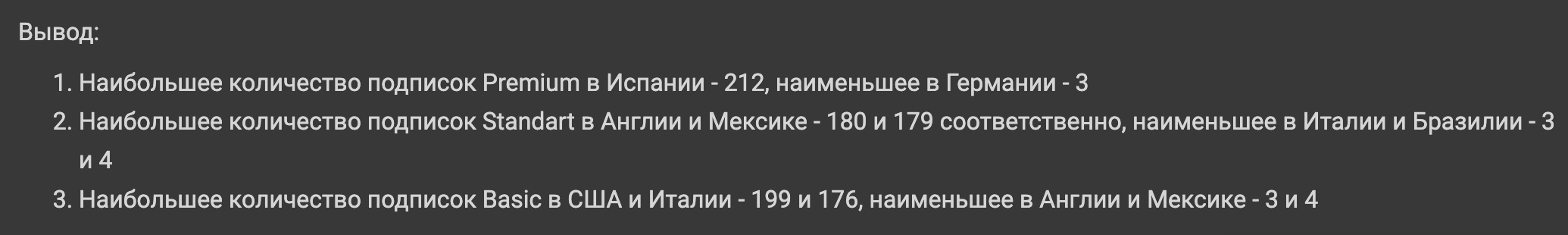




* изучи зависимость типа тарифа от страны пользователя - здесь лучше использовать не корреляцию, а визуализацию: даже если никакой зависимости нет, в любом случае ты сможешь сделать и зафиксировать какие-то выводы.







Когда все задачи будут выполнены, представь общий итог, выделив наиболее важные закономерности и выводы.



