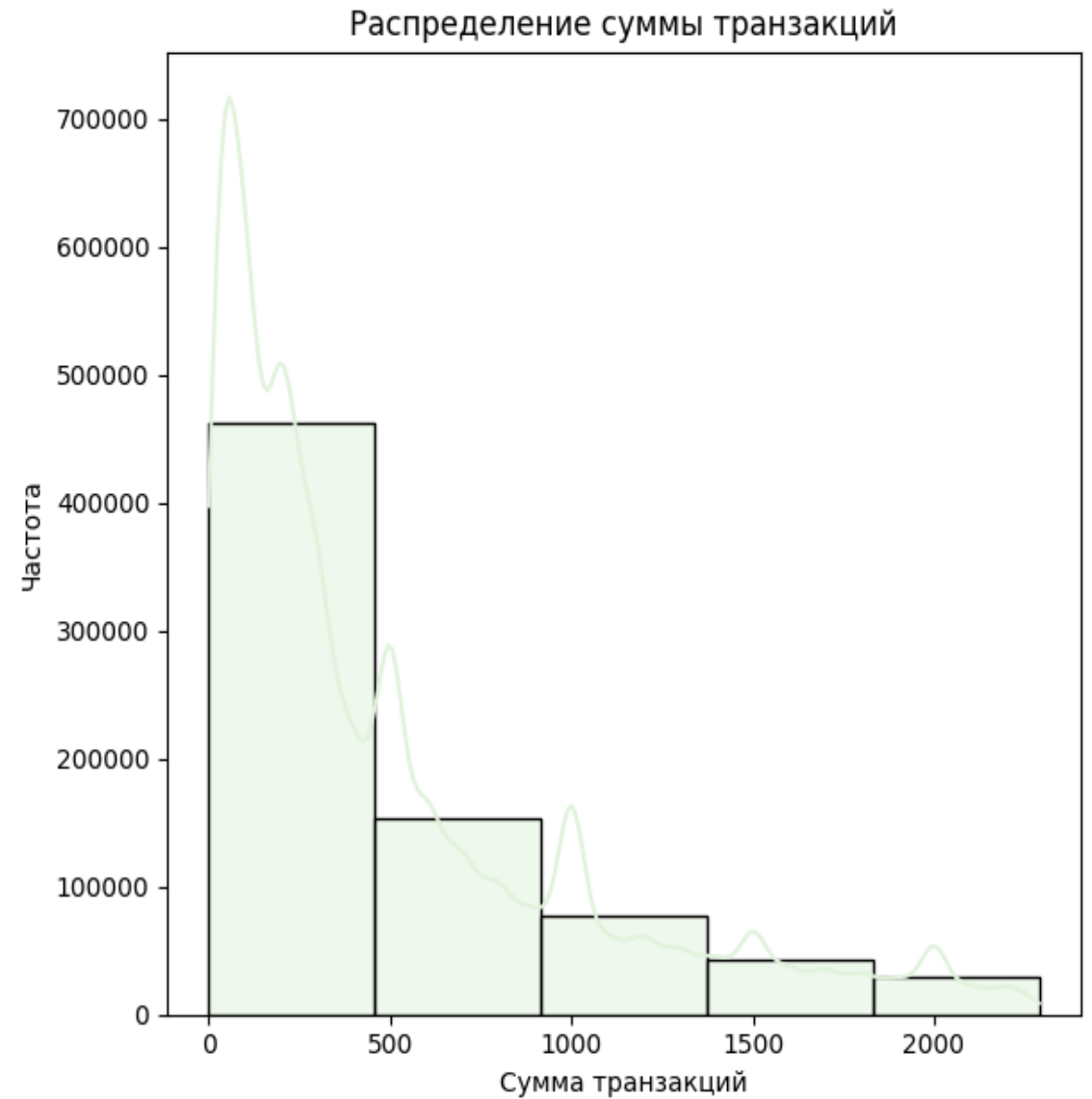
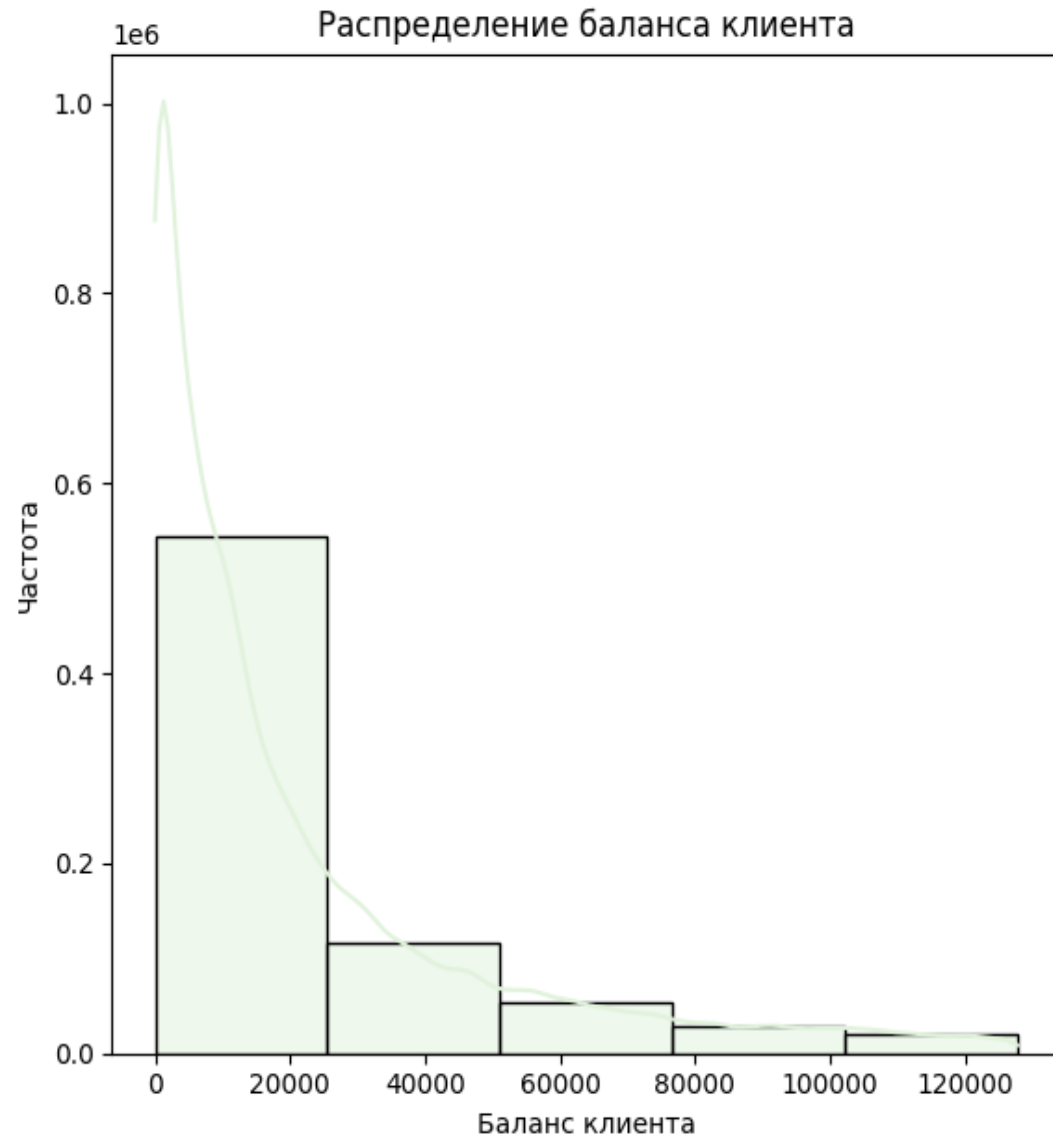


Исследование банковских клиентов

ВАНК

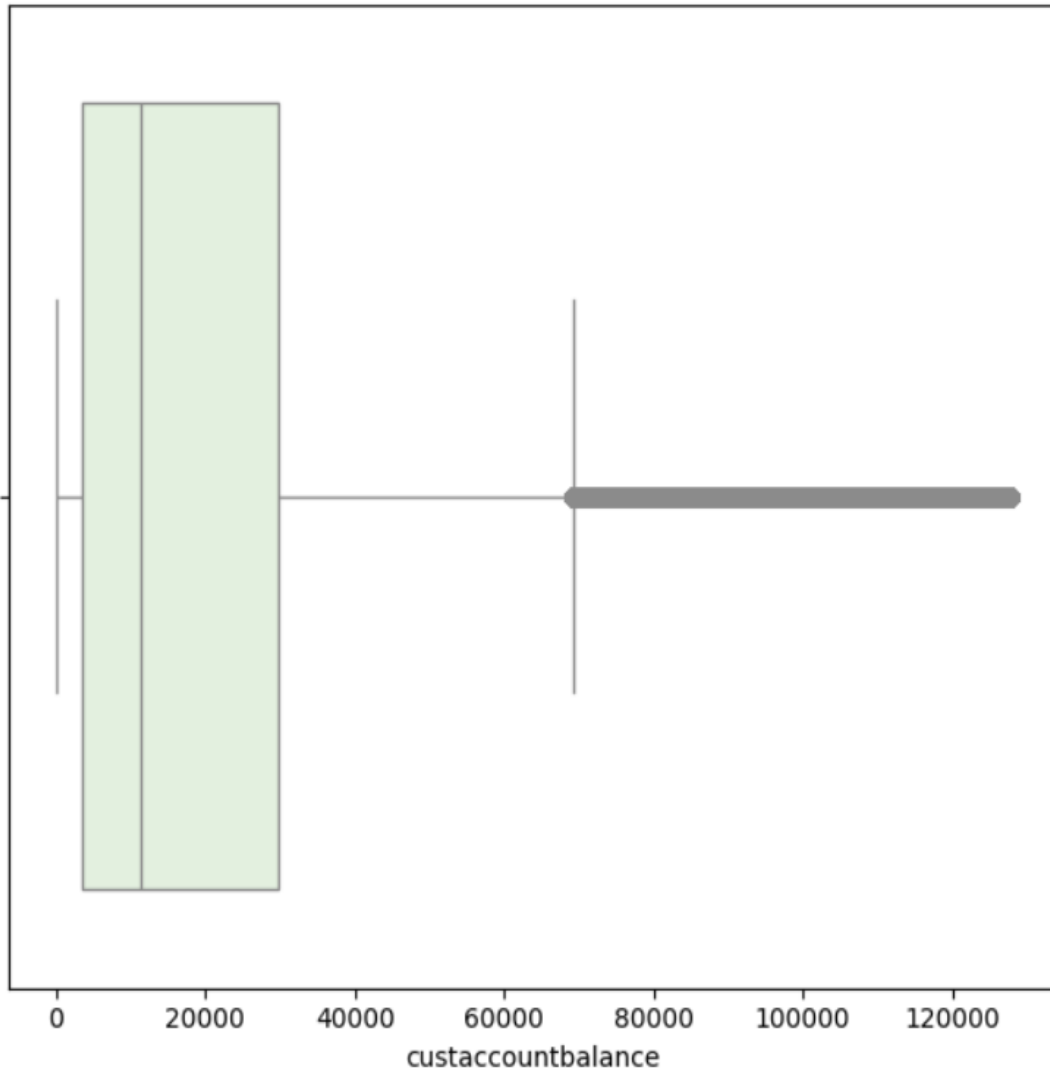
Исследование выполнил Александров
Михаил Александрович в рамках курса по
анализу данных

Распределение баланса клиента и суммы транзакций

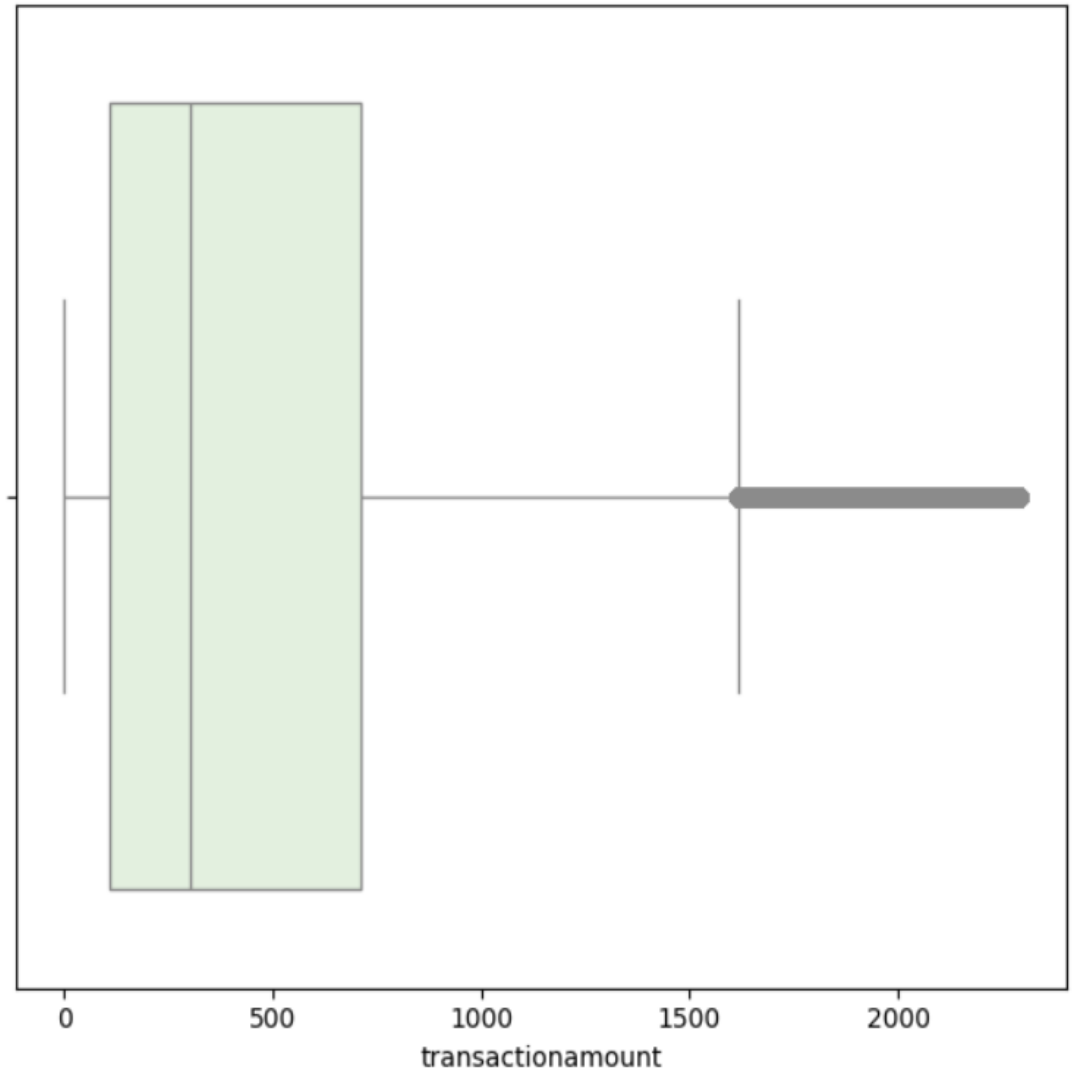


Ящики с усами для баланса клиента и для суммы транзакций

Ящик с усами для баланса клиента

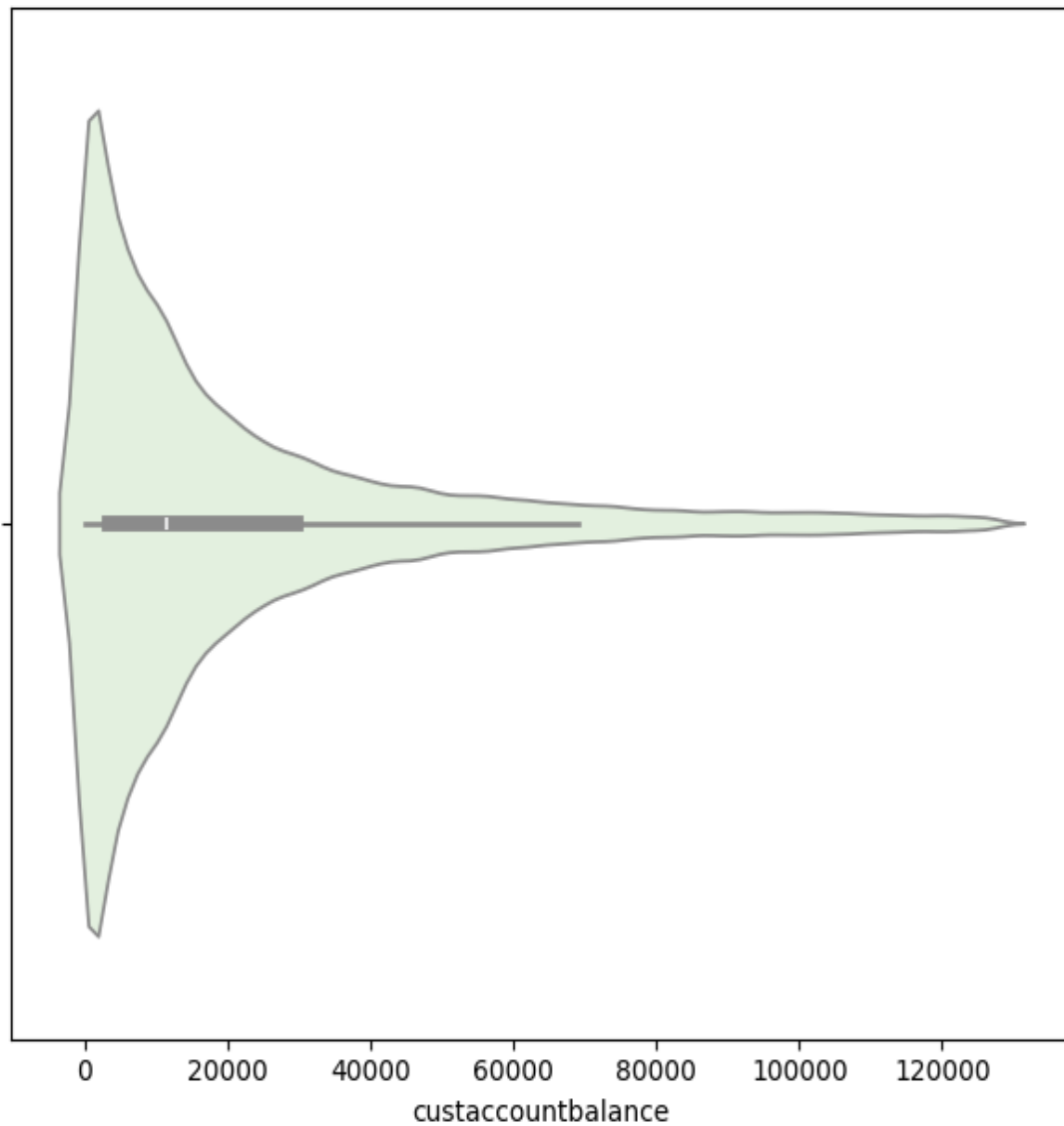


Ящик с усами для суммы транзакций



Скрипичные ключи для баланса клиента и для суммы транзакций

Скрипичный ключ для баланса клиента

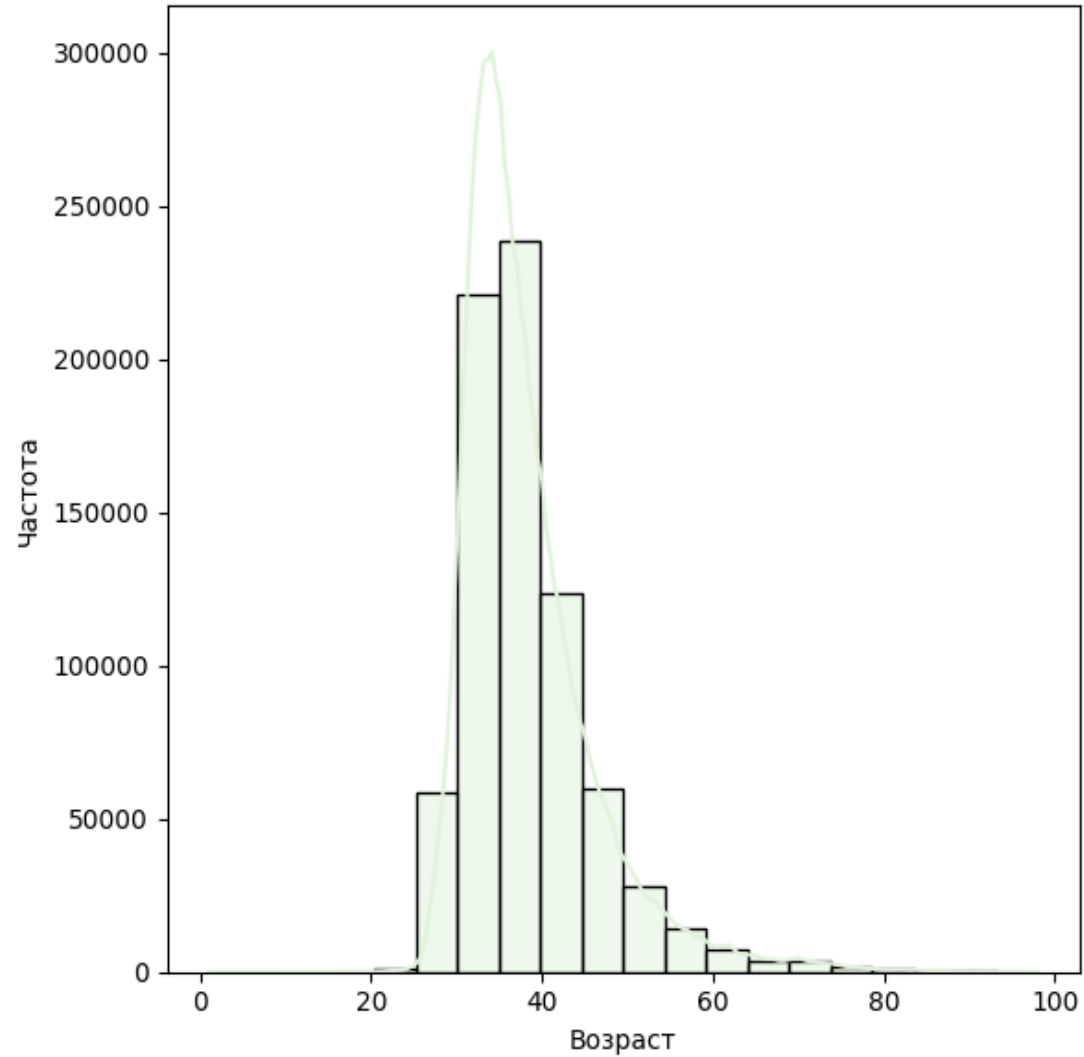


Скрипичный ключ для суммы транзакций

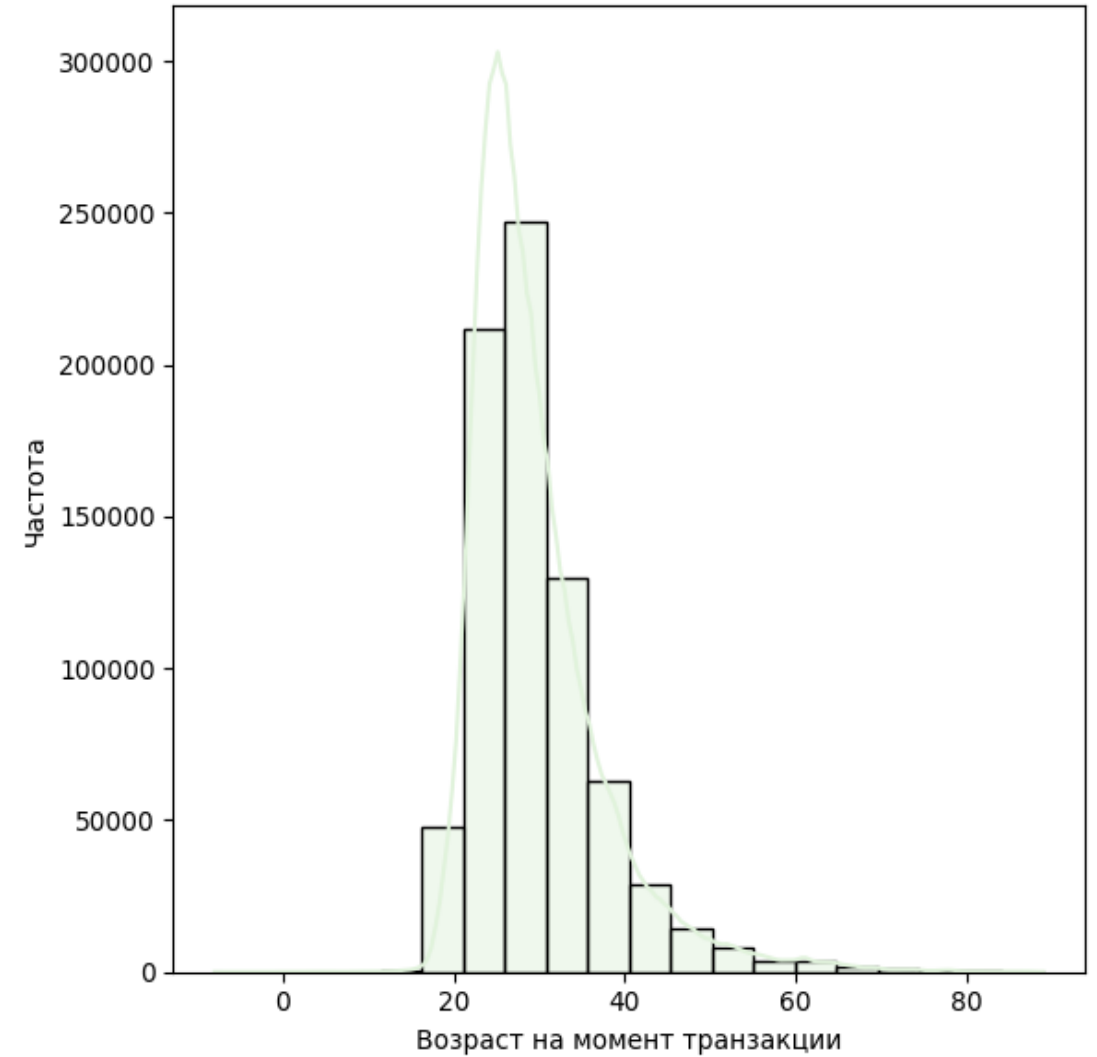


Распределение возраста клиентов

Распределение текущего возраста клиентов

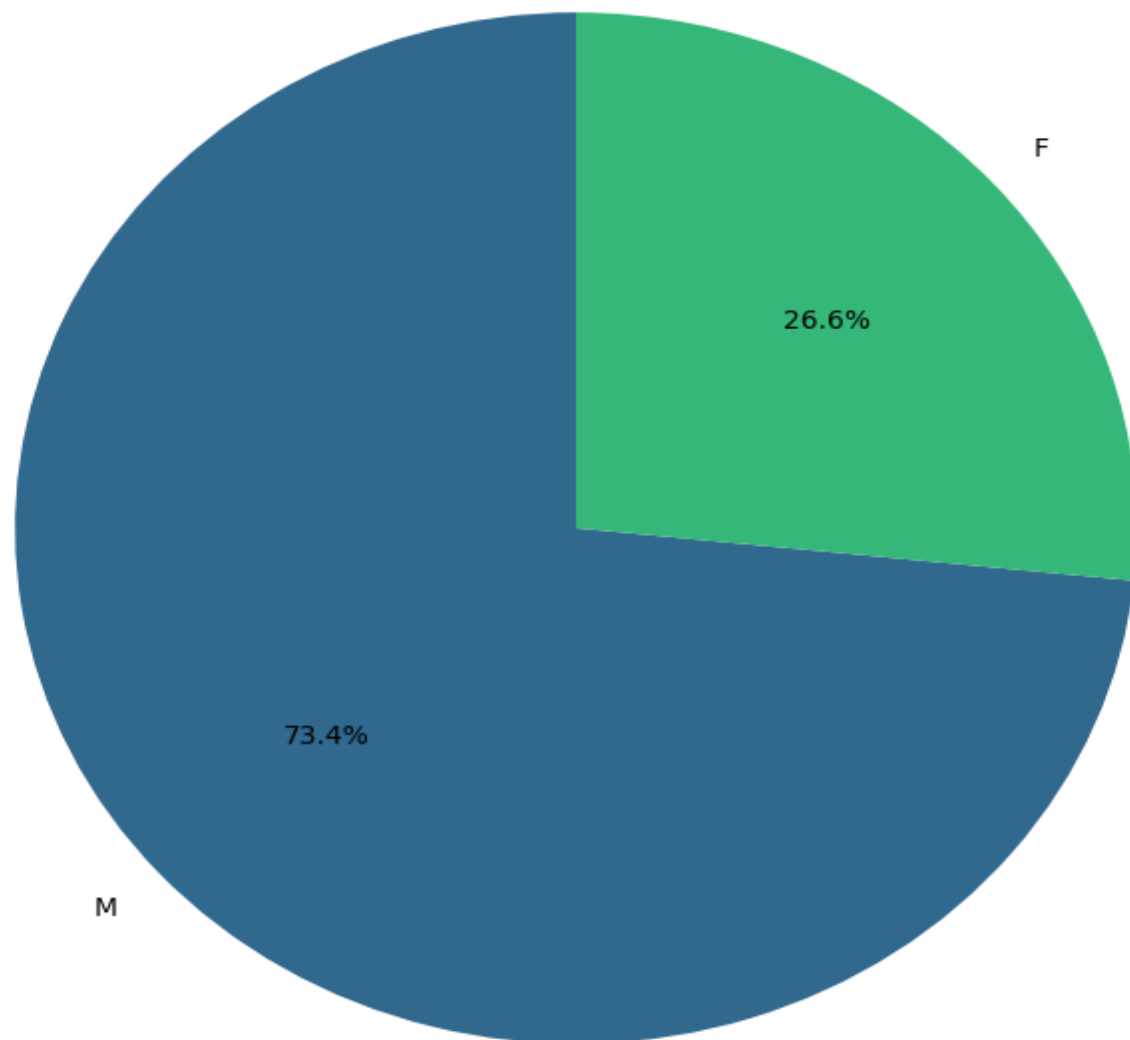


Распределение возраста клиентов на момент транзакции

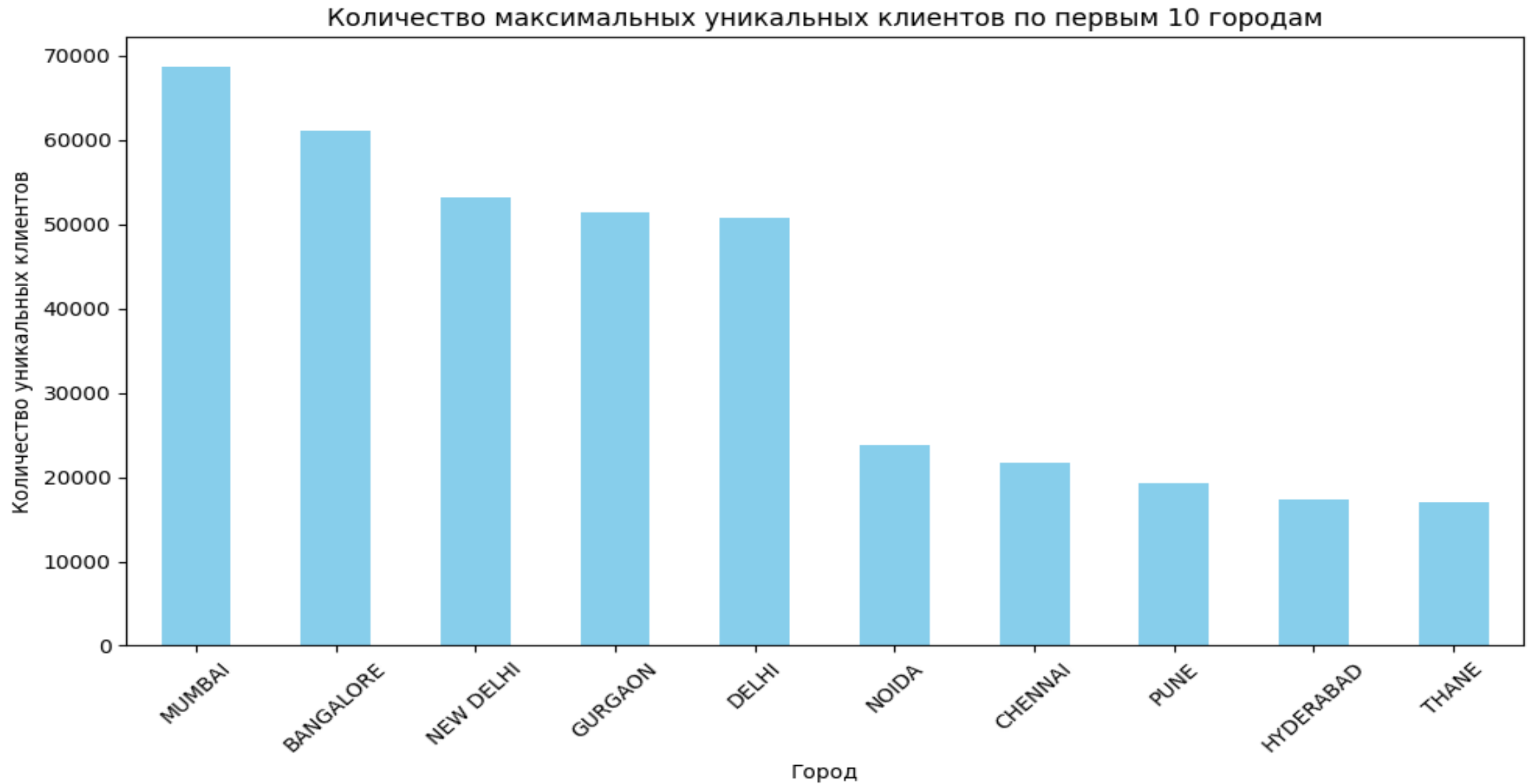


Соотношение клиентов по полу

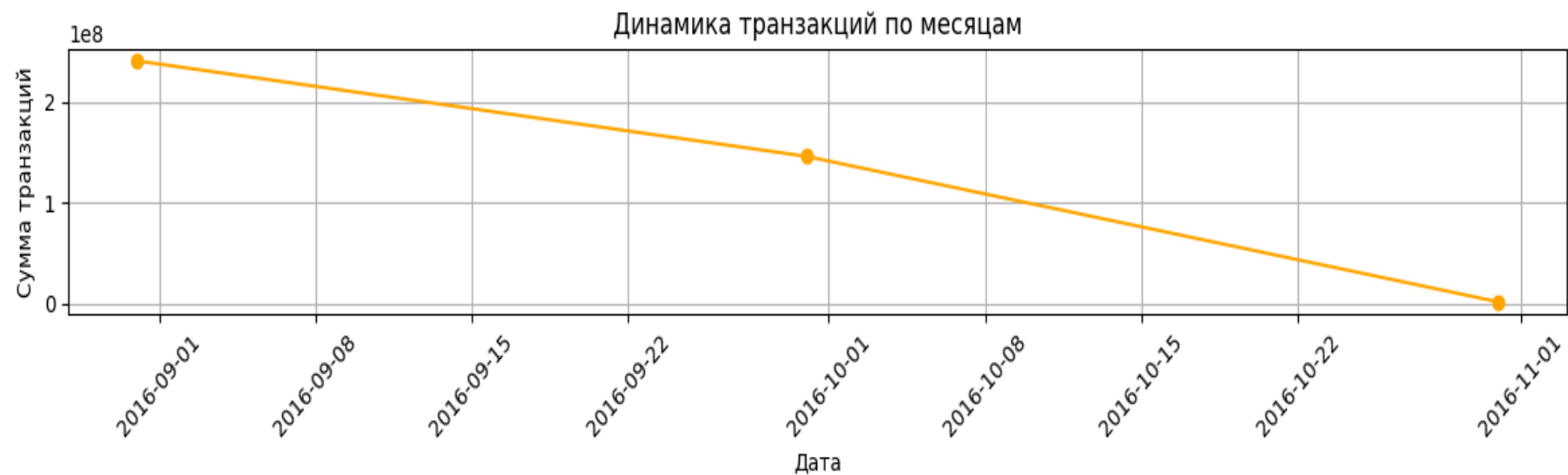
Соотношение клиентов по полу



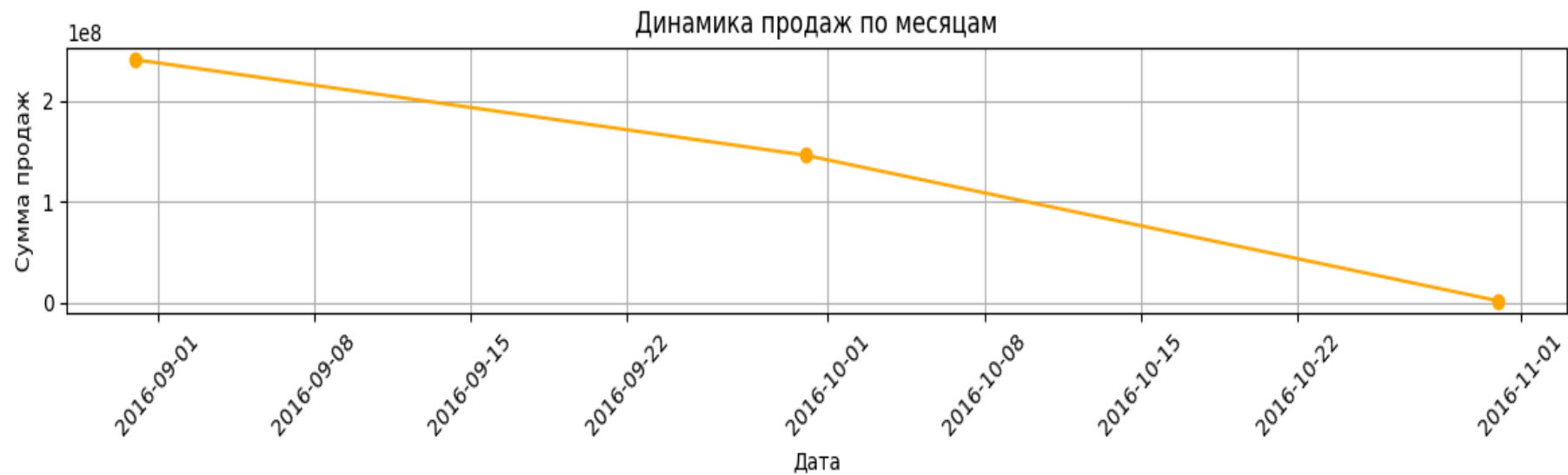
Количество максимальных уникальных клиентов по первым 10 городам



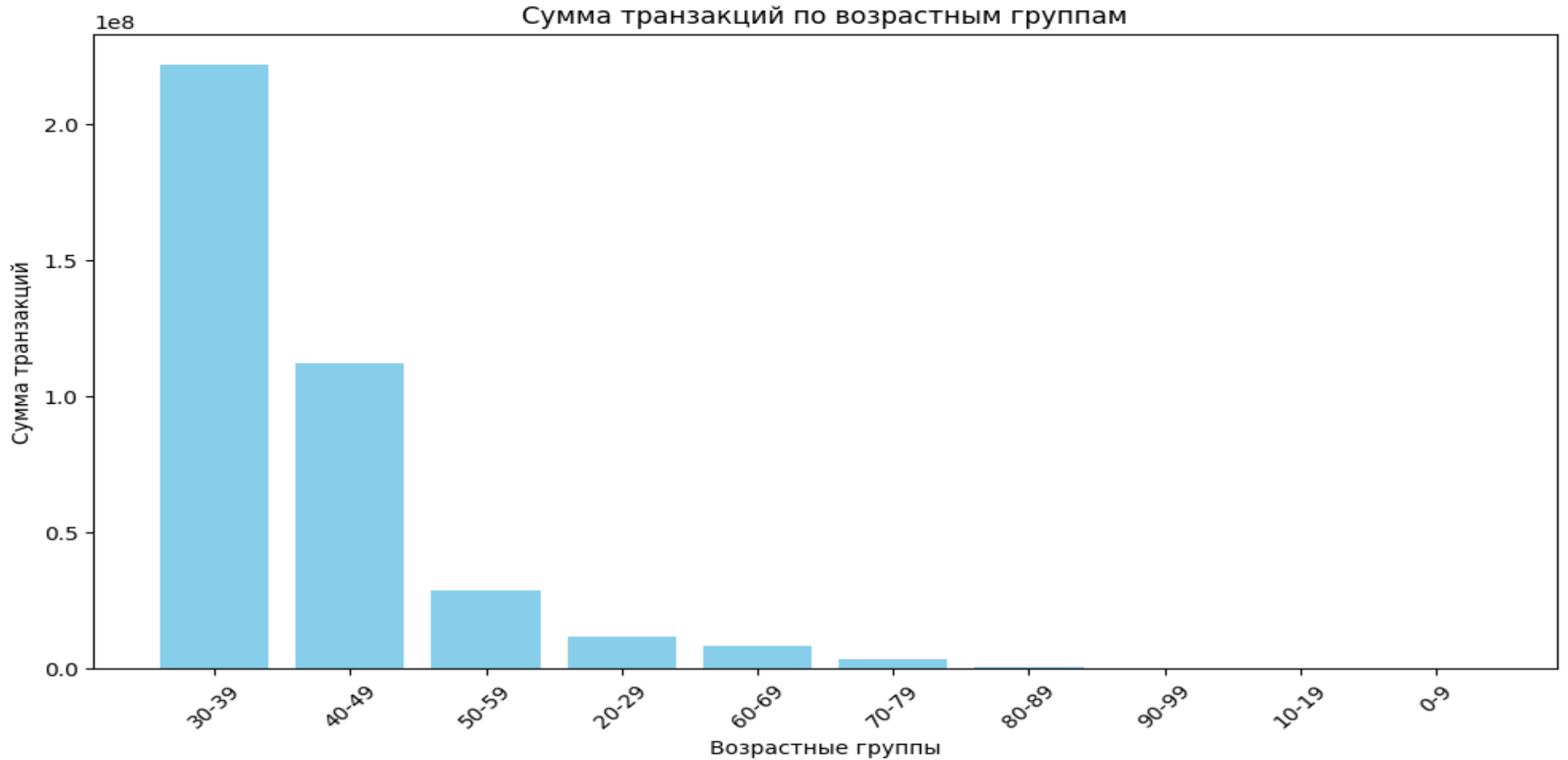
Динамика транзакций по дням и месяцам



Динамика продаж по дням и месяцам

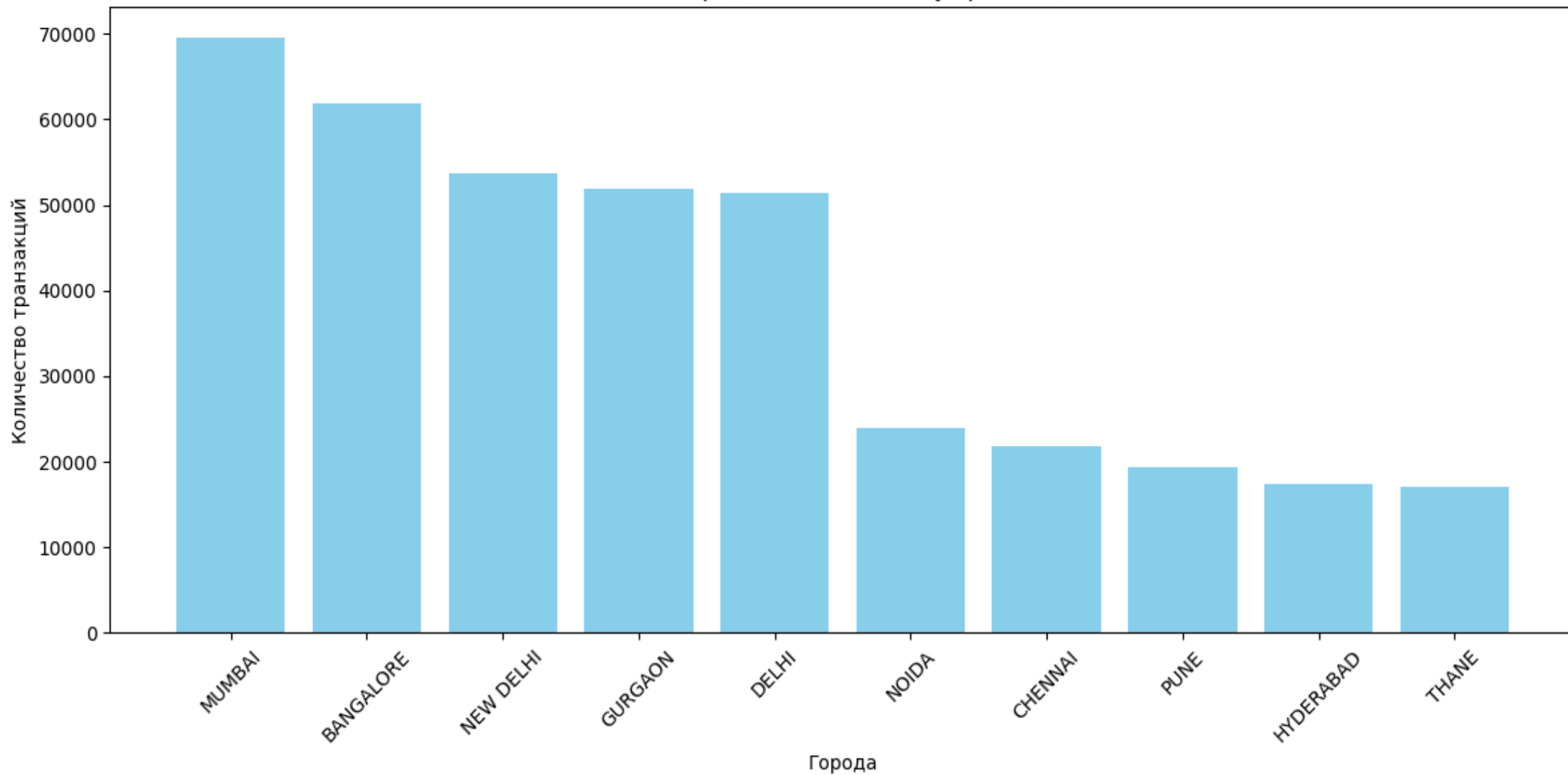


Сумма транзакций по возрастным группам

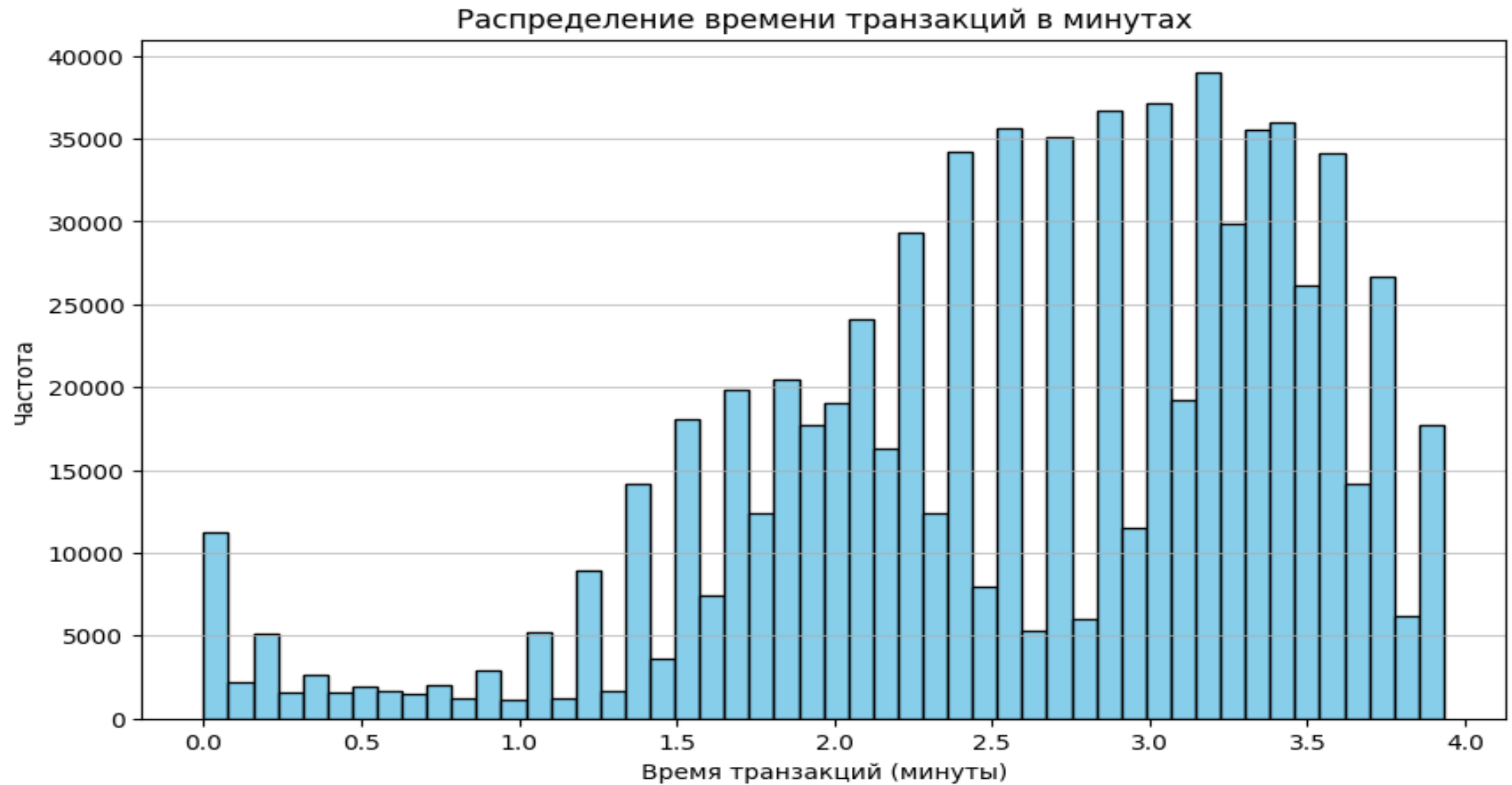


Топ-10 городов по количеству транзакций

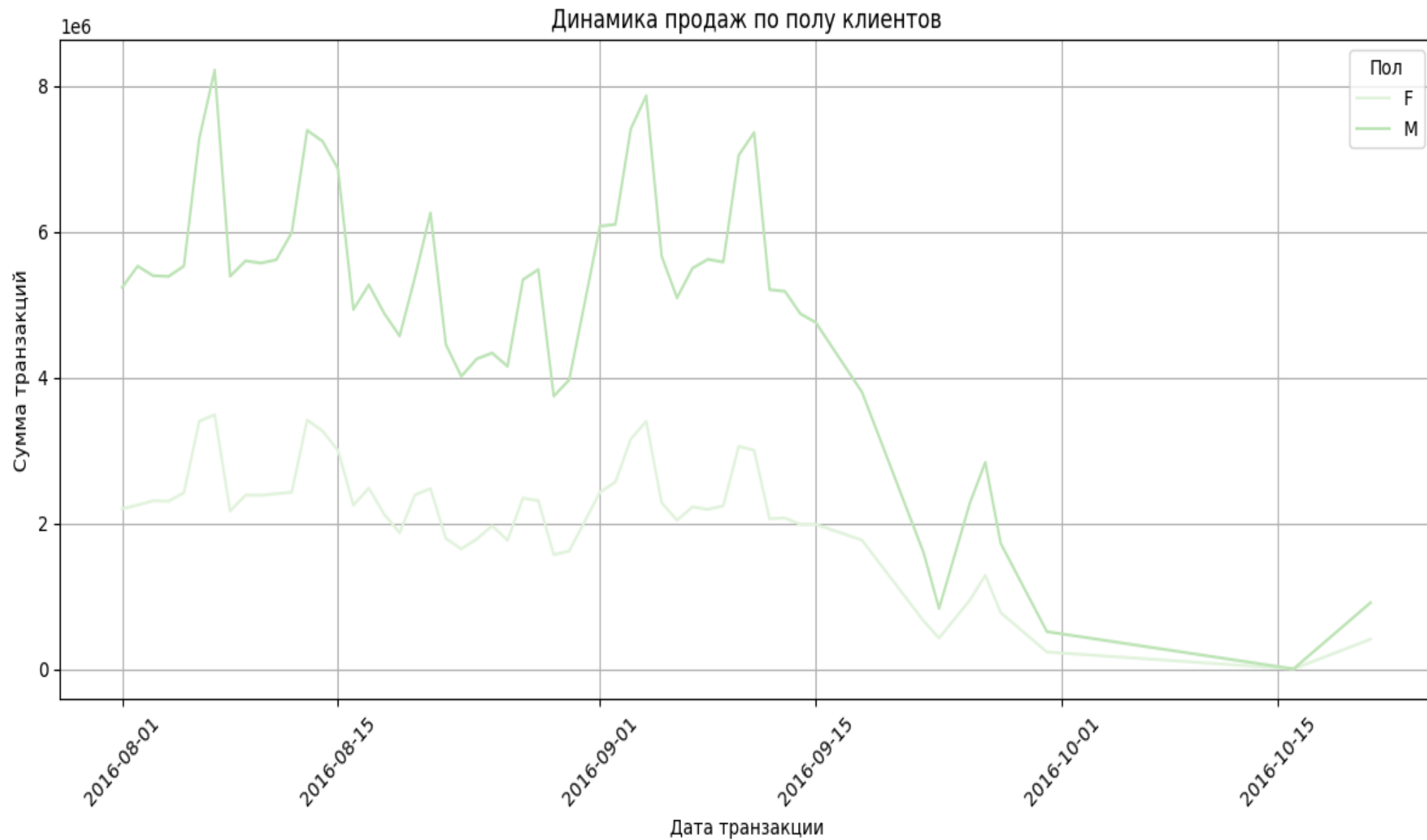
Топ-10 городов по количеству транзакций



Распределение времени транзакций в минутах



Динамика продаж по полу клиентов



Выводы:

1. Загрузили до 1048567 строк данных.
2. Но по некоторым столбцам есть пропуски в данных.
3. Данные в CustomerDOB и CustAccountBalance имеют короткий символьный формат данных даты.
4. TransactionTime представлен в миллисекундах.
5. Данные в customerdob и custAccountbalance приведены в формат данных datetime.
6. transactiontime преобразован в формат минуты:секунды.
7. Дата рождения клиента подкорректирована.
8. Выполнена очистка данных:
9. Удалены дубликаты строк;
10. Удалены строки с пропусками в данных;
11. Получили очищенные данные для проведения аналитики.
12. Среднее количество транзакций на клиента: 1.1340772738894405
13. Гистограммы показывают распределение значений с учётом плотности, что позволяет увидеть, как значения распределены по диапазону: Custaccountbalance от 0.0100 до 127800.120; Custaccountbalance от 0.0100 до 2290.000
14. Ящики с усами в нашем исследовании дают следующие значения о медиане, квартилях и выбросах, что важно для понимания вариации данных: custaccountbalance transactionamount count 762921.000 762921.000 mean 22025.351939 507.886932 min 0.0100 0.0100 25% 3342.300 110.000 50% 11389.6700 300.000 75% 29667.7100 712.3700 max 127800.1200 2290.000 std 27068.588517 522.868333
15. Скрипичные ключи показывают распределение и плотность данных. И на обоих скрипичных ключах визуально прослеживается, что большинство наших клиентов оперируют небольшими суммами.
16. Самый активный возраст наших клиентов составляет от +- 30 лет до 50 лет.
17. Среди наших клиентов: мужчины -73.4%, женщины – 26.6%.
18. Максимальное количество наших клиентов проживают в городе MUMBAI.
19. Максимальное число транзакций наших клиентов наблюдается в возрасте от 30-39 и 40-49 лет.
20. Время транзакций большинства наших клиентов (до почти 40000) составляет 3.25 минуты.

Спасибо за внимание!

BANK