

Анализ оттока клиентов банка «Метанпромбанк»

Исследование провела: Комягина Арина

Дата: декабрь 2022

Источник данных: данные от заказчика, «bank_dataset.csv»

Контакты для связи: mail@mail.ru

Задачи исследования

- Провести исследовательский анализ данных, найти закономерности
 - Составить портрет клиентов, склонных уходить из банка
 - Проверить гипотезы о равенстве дохода между ушедшими и оставшимися клиентами и о равенстве баллов кредитного скоринга между ними.

Выводы

Между множествами ушедшими и оставшимися клиентами есть заметные различия в:

- Возрасте
- Количестве используемых ими продуктов
- Активности
- Пола
- Географии

Портрет склонных уходить клиентов:

- Женщины
- Старше 45
- Из Ростова Великого
- Использовал(а) больше 2 продуктов
- Не проявлял(а) активность в последнее время

Портрет не склонных уходить клиентов:

- Мужчины
- От 18 до 25 лет
- Из Ярославля
- Использовал(а) два продукта
- Последнее время проявлял(а) активность

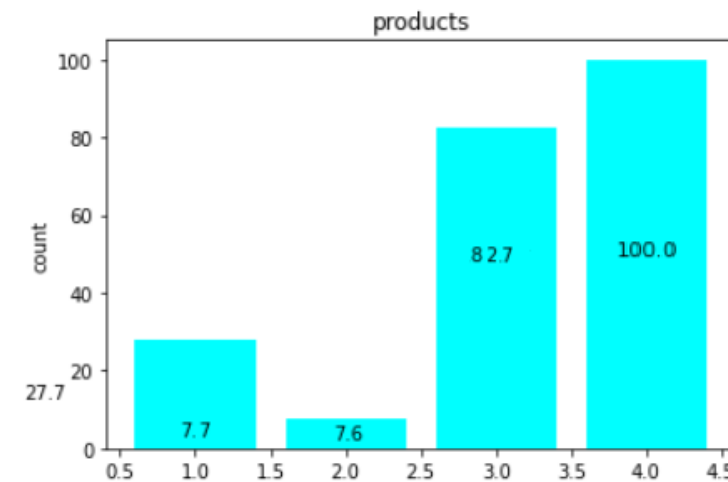
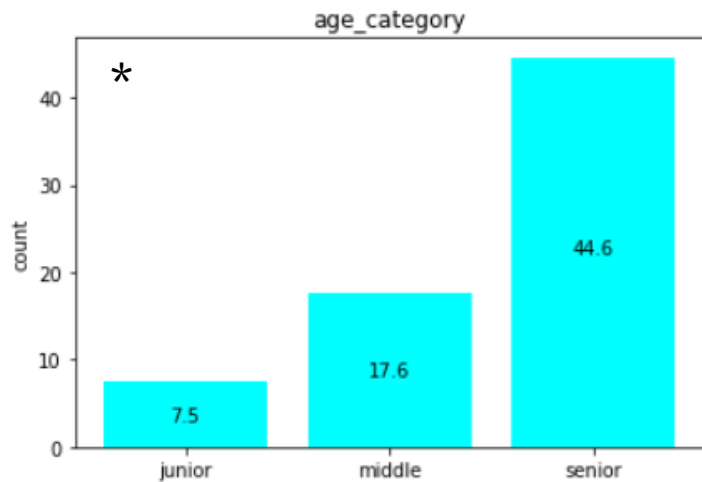
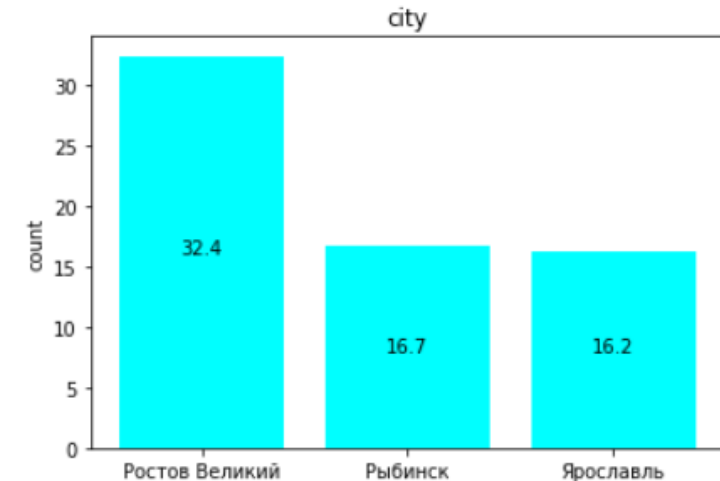
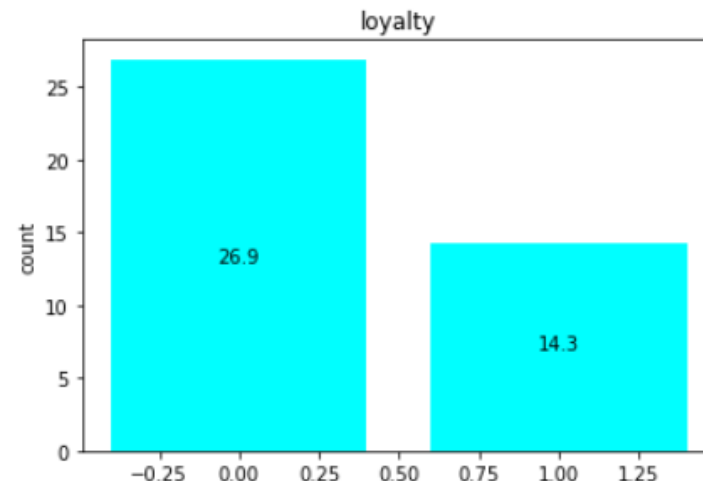
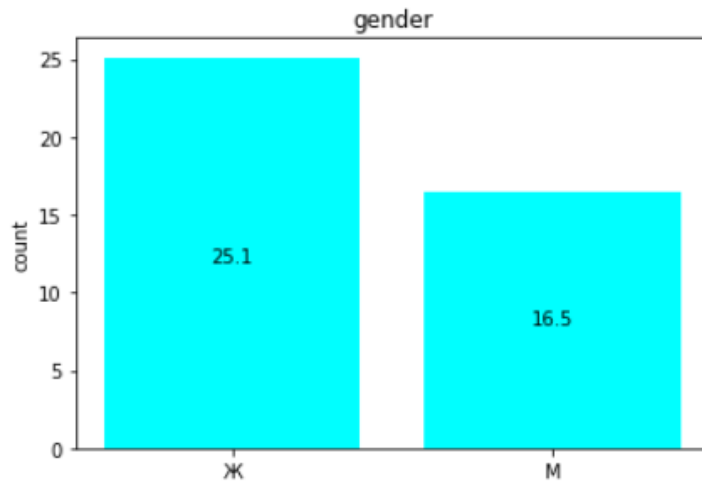
Исследование гипотез:

Выявлено, что между доходом клиентов, ушедших из банка и клиентов, оставшихся в нём, нет статистически значимой разницы, а между их кол-вом баллов кредитного скоринга статистически значимая разница есть.

Рекомендации

1. Проверить данные для столбца balance и уточнить причину пропусков.
2. Разработать продукт, который мог бы заинтересовать уходящих клиентов, дополнительно прорекламировать его. Например, льготные условия вкладов для пенсионеров, обслуживание пенсионного счёта.
3. Изучить, какими продуктами пользовались клиенты, у которых было более двух продуктов. Возможно, в них есть проблема, и это отпугнуло часть клиентов.

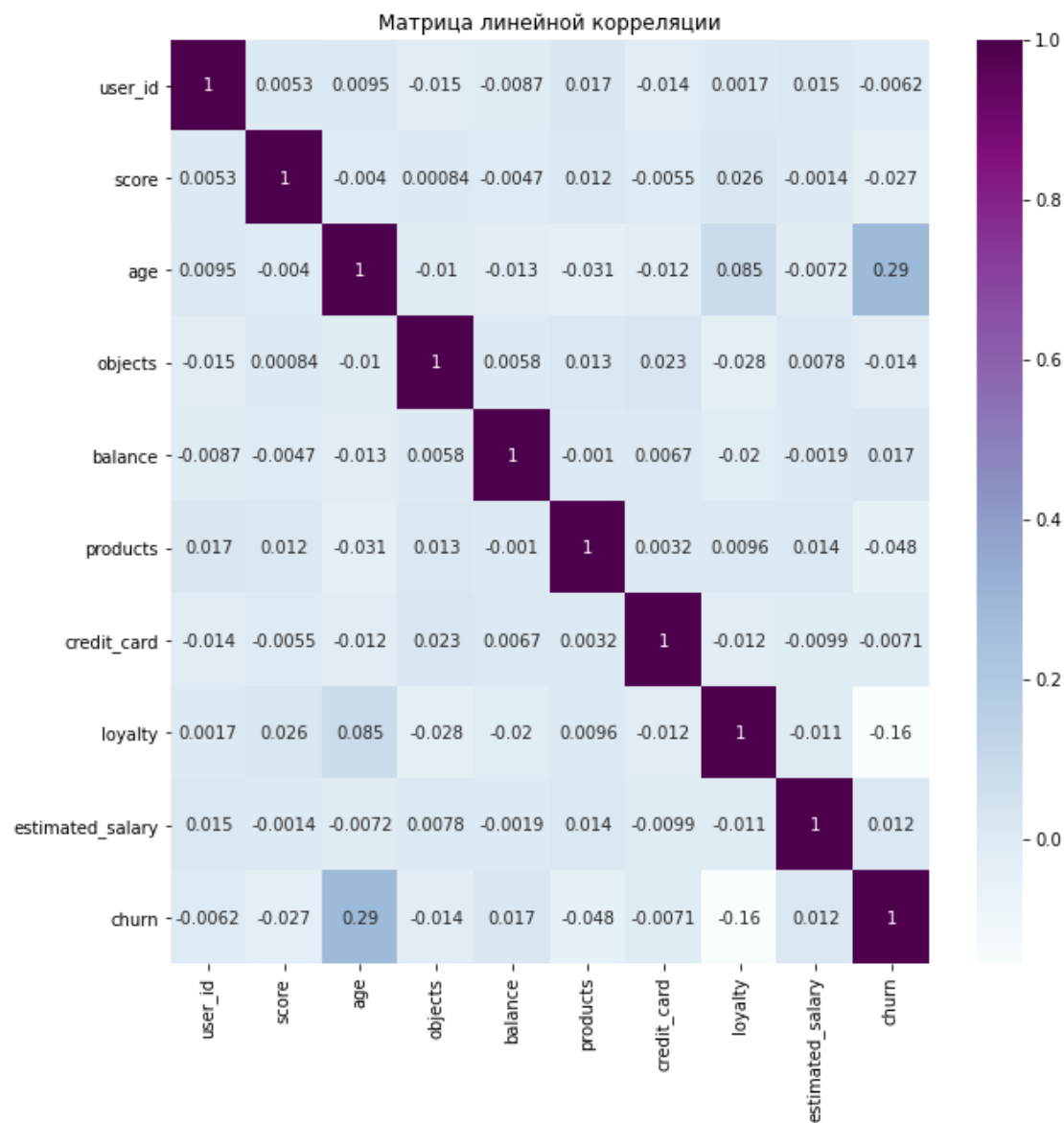
Графики распределения по различным признакам, % ушедших клиентов от общего числа



* Клиентам от 18 (минимальный возраст) до 25 лет присвоена категория junior, от 26 до 45 лет - middle, более 45 - senior.

- Возраст - ушло 44.6% клиентов старше 45 лет, 17.6% - от 25 до 45 лет, и всего 7.5% - от 18 до 25 лет
- Количество используемых продуктов - ушли все клиенты, использующие 4 продукта, 83%- 3 продукта, 8% - 2 продукта и 28% - 1 продукт
- Активность - ушло 26.9% неактивных клиентов и 14.3% - активных
- Пол - ушло 25.1% женщин и 16.5% мужчин)
- География - в Ростове Великом ушло 32,4% клиентов, в Рыбинске - 16.7, в Ярославле - 16.2%

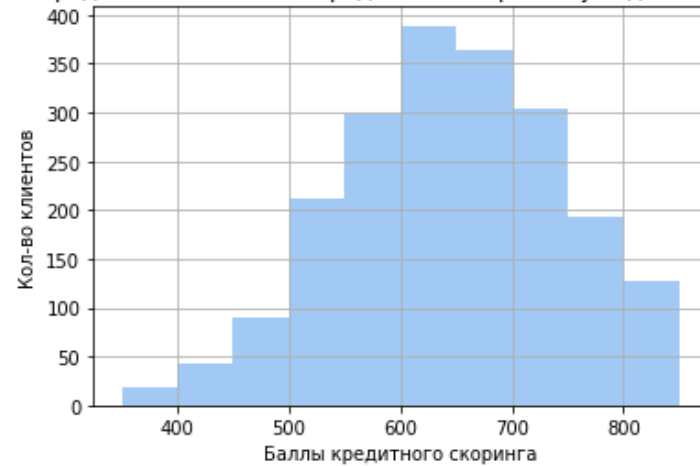
Другие корреляции



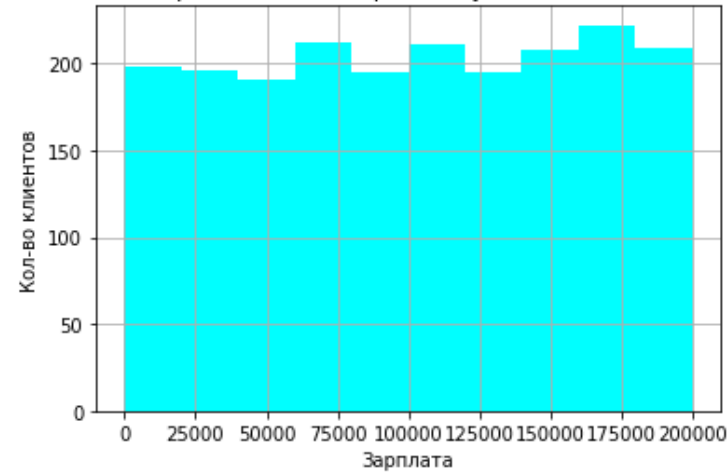
Коэффициент корреляции везде < 0.3 - связь очень слабая. Однако можно отметить связь churn и age, а также loyalty и age - их парный коэффициент корреляции больше остальных и достигает по меньшей мере десятых долей.

Проверка гипотез

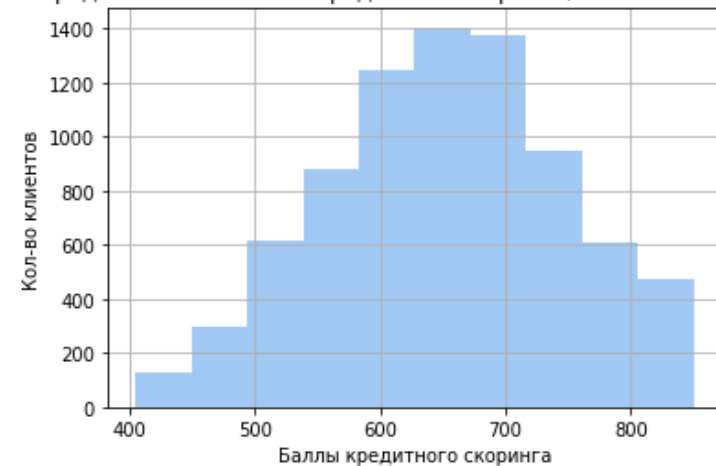
Распределение по баллам кредитного скоринга, ушедшие клиенты



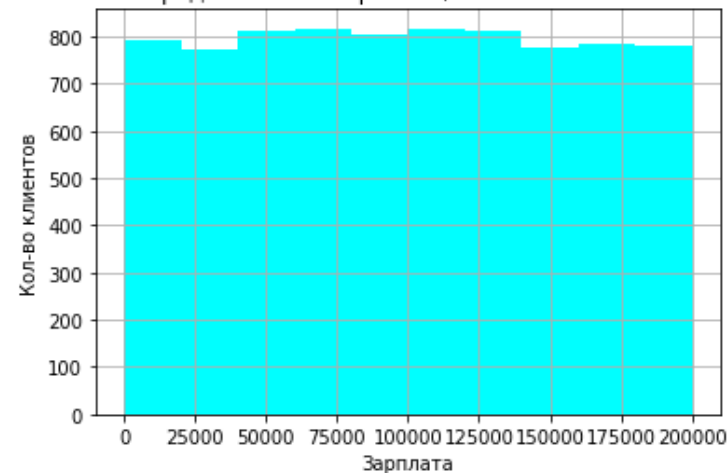
Распределение по зарплате, ушедшие клиенты



Распределение по баллам кредитного скоринга, оставшиеся клиенты



Распределение по зарплате, оставшиеся клиенты



p-значение: 0.01986866117272834

Отвергаем нулевую гипотезу: разница статистически значима

p-значение: 0.22705150930153783

Не получилось отвергнуть нулевую гипотезу, вывод о различии сделать нельзя