

Финальный проект: Банки — Анализ оттока клиентов.

Датасет содержит данные о клиентах банка «Метанпром». Банк располагается в Ярославле и областных городах: Ростов Великий и Рыбинск.

Работу выполнил: Пономарев Д.Д. Da_67_01
Email: Ponomarev.D.D@yandex.ru

Цели и задачи:

Нашей главной задачей станет анализ оттока клиентов. Анализ покажет, какие клиенты уходят из банка, а так же поможет нам составить сегменты клиентов, которые склонны уходить из банка. «Метанпромбанк» — деньги не пахнут!

Проанализируем клиентов регионального банка и выделим сегменты клиентов, которые склонны уходить из банка.

Проведём исследовательский анализ данных, определим все значимые признаки отточности (интервалы значений характеристик, которые связаны с повышенным оттоком, сравните портреты типичных клиентов, которые склонны и не склонны уходить из банка и т.д)

Сформулируем и проверим статистические гипотезы.

Проверим гипотезу различия дохода между теми клиентами, которые ушли и теми, которые остались.

Проверим гипотезу различия баллов кредитного скорринга между теми клиентами, которые ушли и теми, которые остались.

Объединяя признаки отточности, сформируем сегменты, отберем из них лучшие и дадим по ним рекомендации

ОБЩИЙ ВЫВОД:

в исследовательском анализе мы составили портреты отточных и неотточных клиентов и сравнили их:

У отточных выше средний балл собственности на 58%.

Средний баланс на 55%.

Среднее количество продуктов на 35%.

Средняя активность на 45%.

Доля мужчин у отточных выше на 38%.

Средний возраст отличается незначительно - на 3.6%.

Средний кредитный скоринг - на 2%.

Построили гистограммы распределения непрерывных в разрезе оттока:

для баллов кредитного скоринга можно выделить отточные интервалы с 840 до 950

для возраста с 25 до 36, с 49 до 61.

для баланса виден повышенный отток в интервалах с 1100000 и выше.

для заработной платы от 100000 до 140000, со 155000 до 210000.

Визуализировали для категориальных переменных доли оттока:

у городов чуть выше среднего отток у Ярославл и Ростова. У рыбинска ниже среднего.

По гендерному признаку более склонны к оттоку мужчины, у женщин отток ниже среднего.

Клиенты с кредитной карты уходят реже.

Активные клиенты чаще склонны к оттоку.

По баллам собственности клиенты с 1-3 склонны к оттоку ниже среднего, с 4-9 выше. Клиенты с 9 продуктами уходят в 50% случаев. У клиентов с 1 продуктом вероятность оттока ниже средней, с 2-мя примерно равна средней, с 3-5 выше. У клиентов с 4 продуктами % оттока выше 60%!

Построена матрицу корреляций ϕ_{hik} :

связь слабая для: индентификатора пользователя, баллов кредитного скоринга, пола, возраста, наличия кредитной карты, активности пользователя.

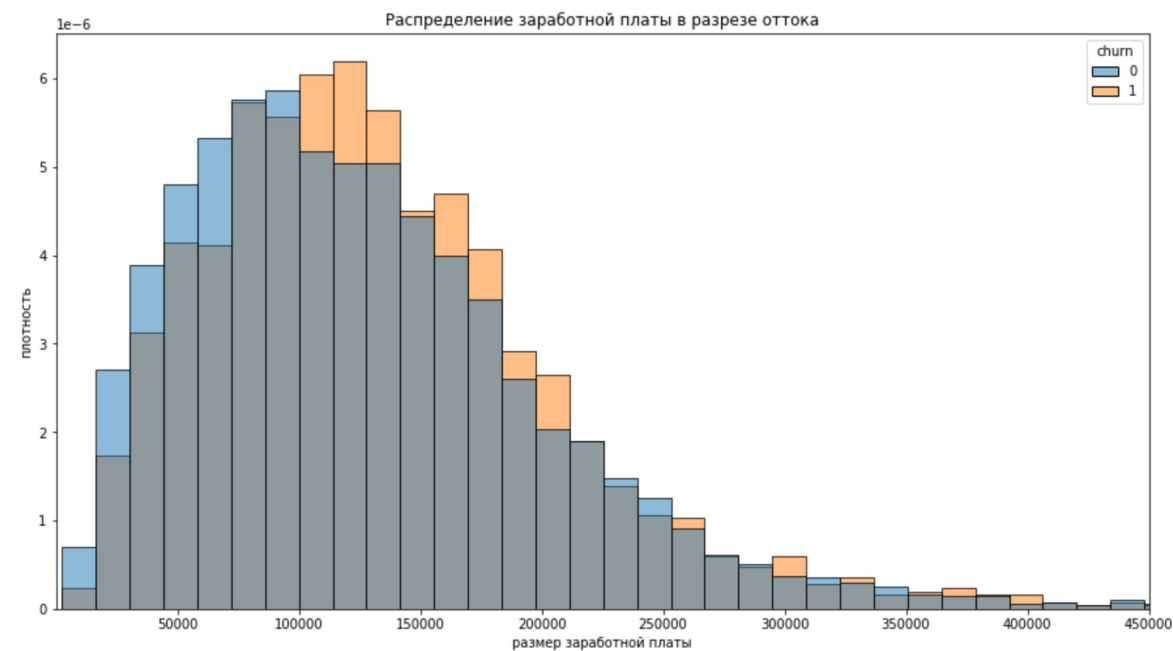
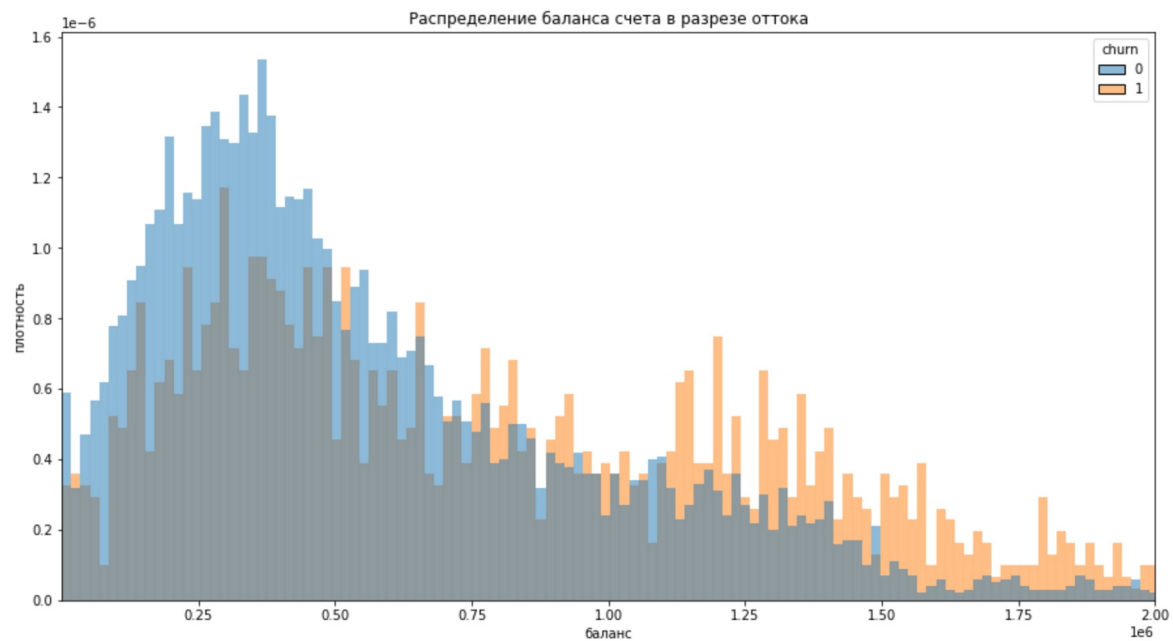
умеренная для: баллов собственности, количества используемых продуктов.

Проверены 2 гипотезы:

существуют статистически значимые различия в доходе между клиентами, которые ушли и которые остались.

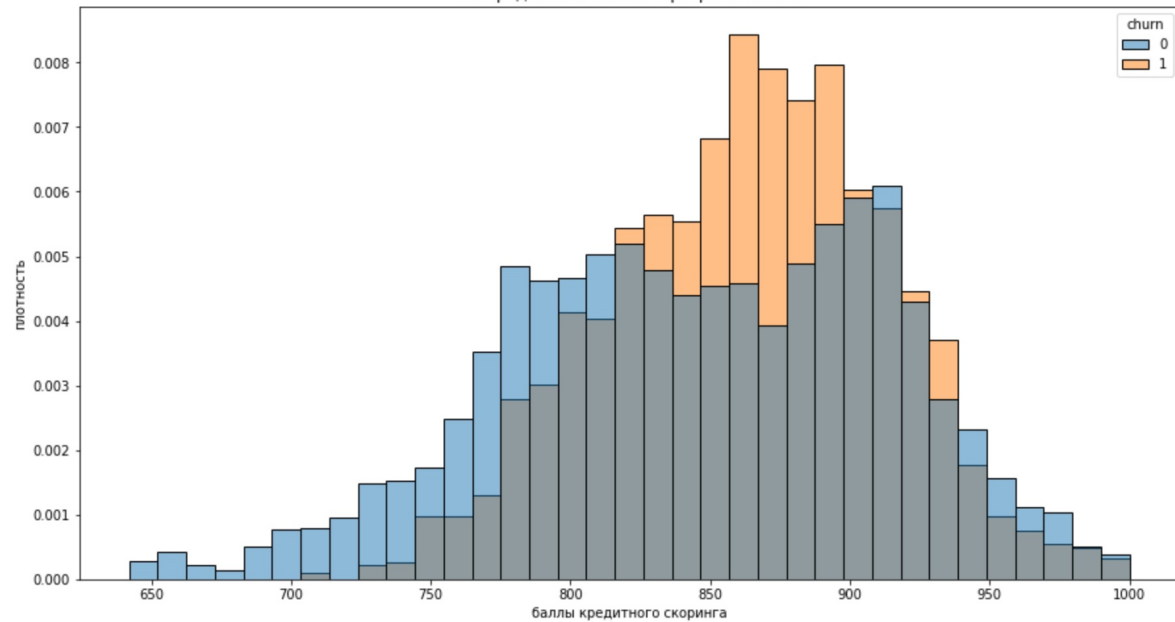
существуют статистически значимые различия в в баллах кредитного скоринга между клиентами, которые ушли и которые остались.

Распределения баланса и заработной платы в разрезе оттока:

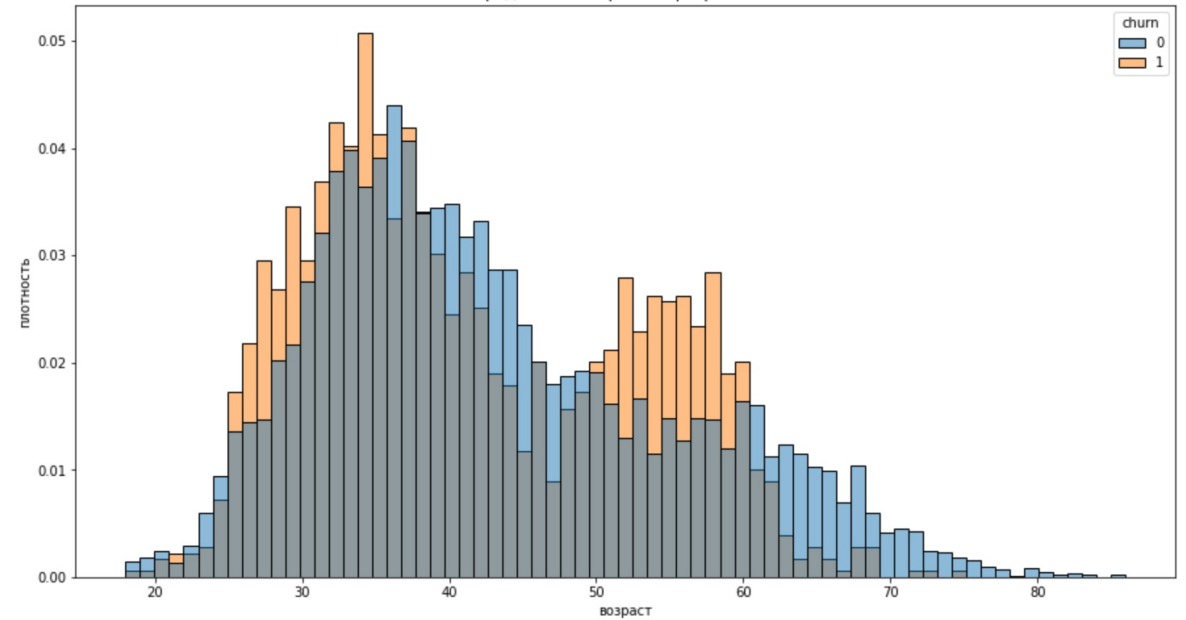


Распределение баллов кредитного скорринга и возраста в разрезе оттока:

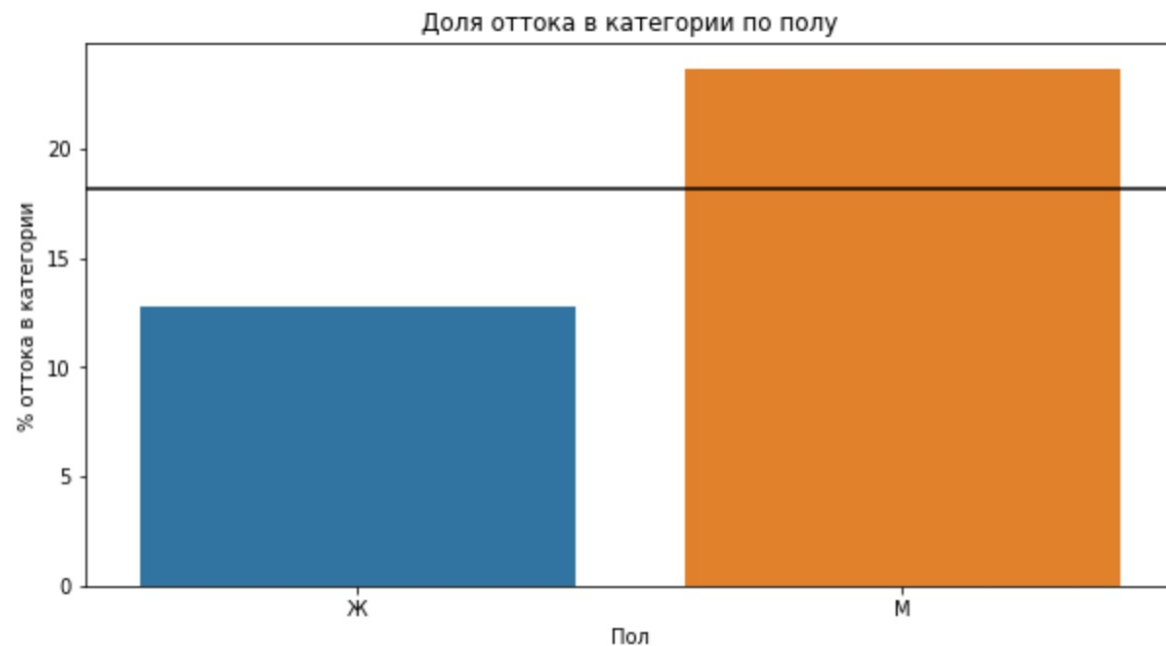
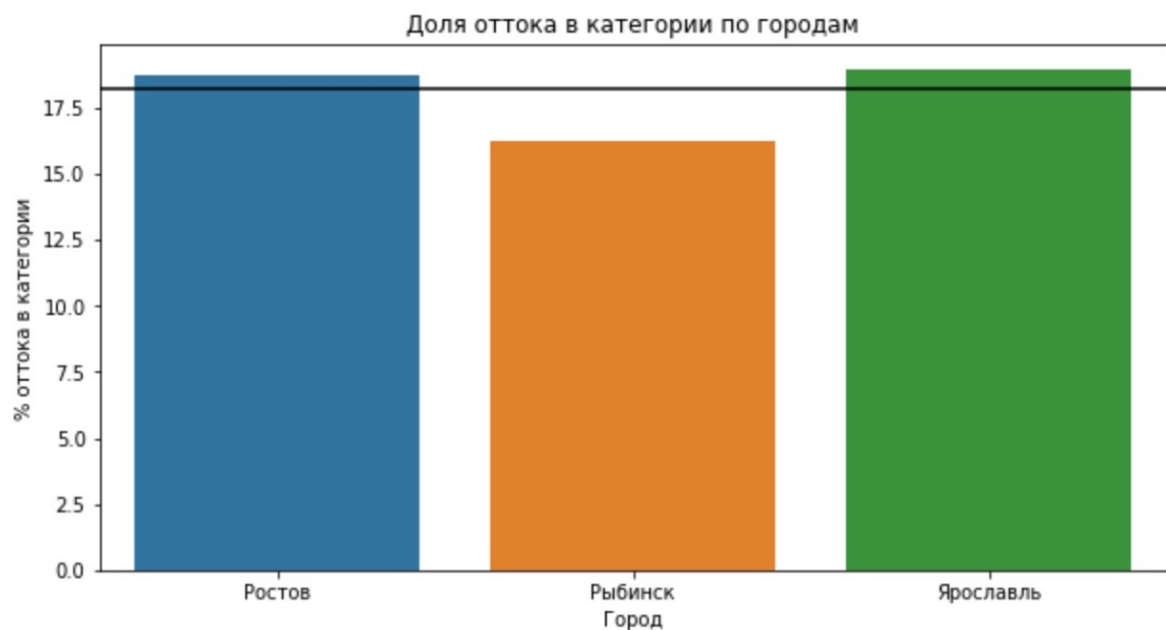
Распределение баллов в разрезе оттока



Распределение возраста в разрезе оттока

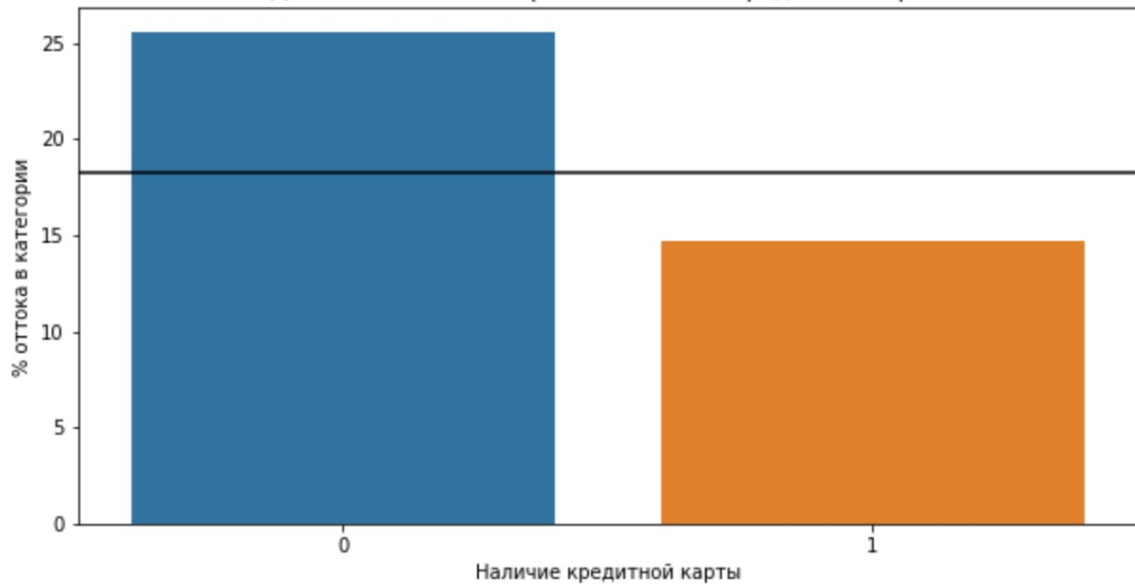


Уровень оттока в категориях по городам и гендерной принадлежности:

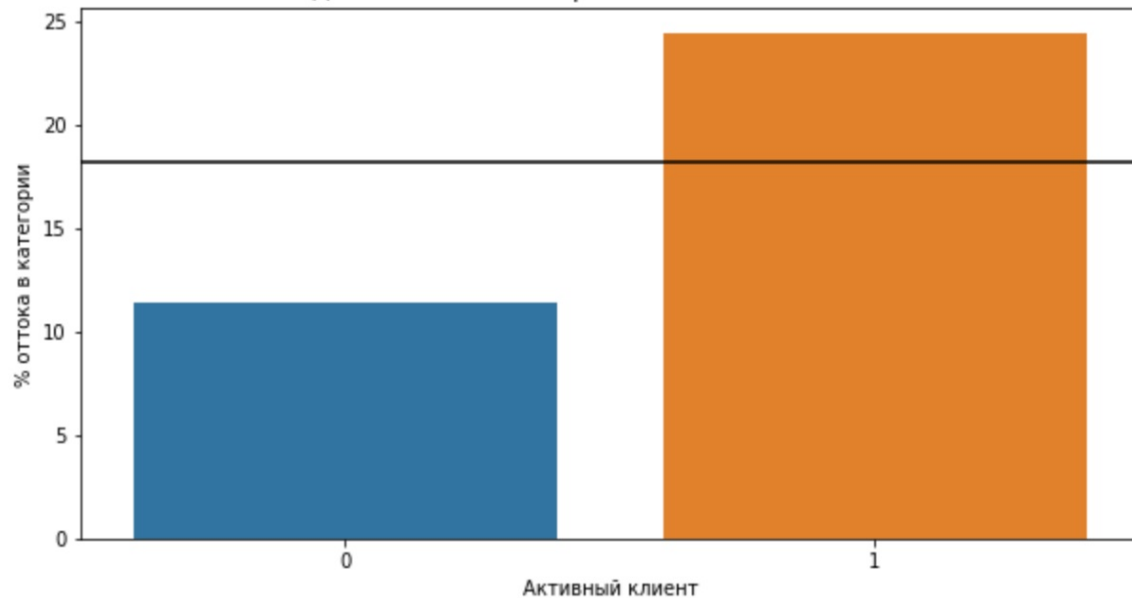


Уровень оттока по наличию кредитной карты и активности клиента:

Доля оттока в категории по наличию кредитной карты

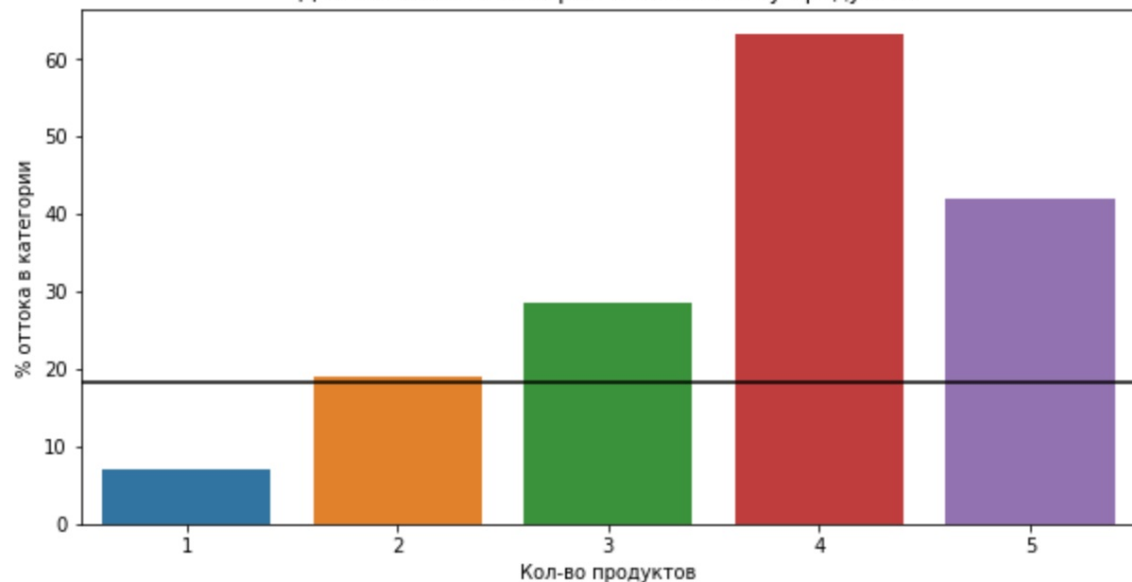


Доля оттока в категории по активности клиента



Уровень оттока по количеству используемых продуктов и баллов собственности:

Доля оттока в категории по количеству продуктов



Доля оттока в категории по баллам собственности

