

# Анализ оттока клиентов банк "Метанпром"

Выпускной проект

Студент: Ни И.Р

Группа: DA-35

# Проблема

## Ситуация

В банке наблюдается значительное увеличение оттока клиентов, что угрожает его стабильности.

## Контекст

Заказчик - Отдел маркетинга банка "Метанпром"

Ранее проводились исследования, но они устарели и не отражают текущую ситуацию.

## Задача


Максимально эффективно выделить сегменты клиентов, склонных к оттоку, чтобы они были однородными и покрывали максимум проблемных зон банка

# Цель и задачи

## Цель проекта:

- Выявление сегментов клиентов с высоким риском оттока

## Основные задачи:

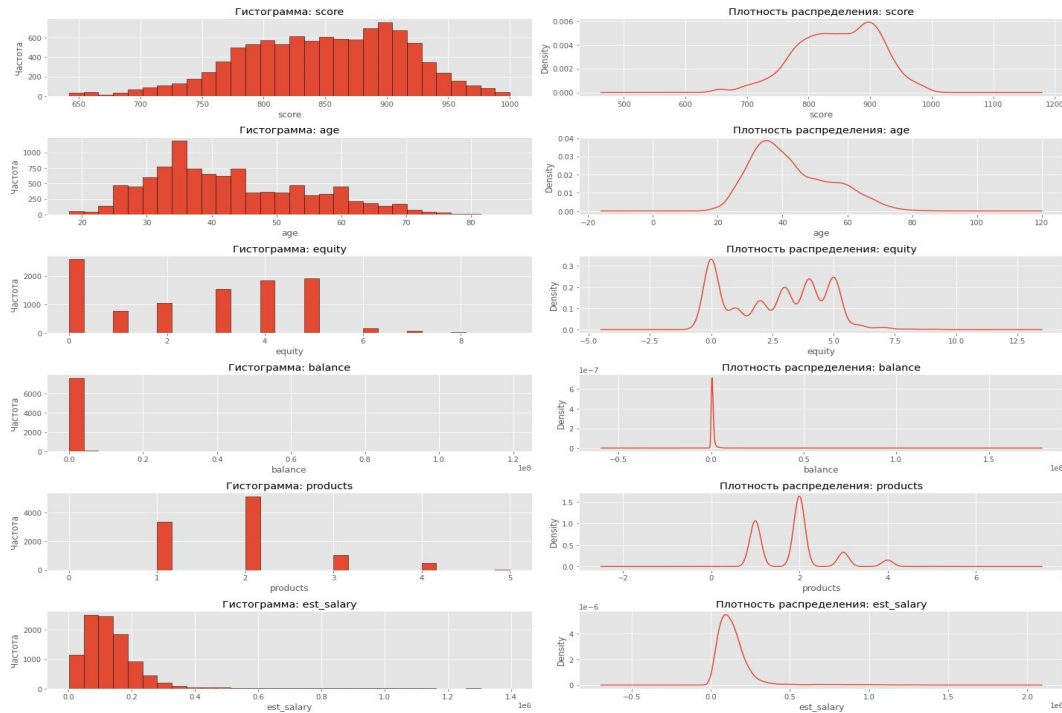
1. Проведение исследовательского анализа данных (EDA).
  2. Выделение ключевых характеристик, влияющих на отток.
  3. Проверка статистических гипотез о факторах, влияющих на отток.
  4. Формирование рекомендаций по сегментации и маркетинговым воздействиям.
- 

# Общая информация:

Количество клиентов в выборке: 10 тыс.

# Числовые параметры

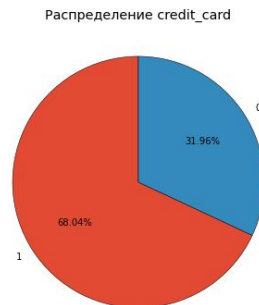
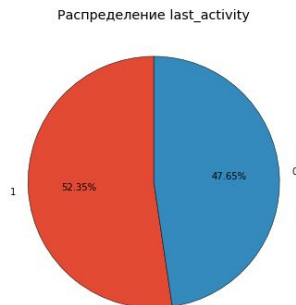
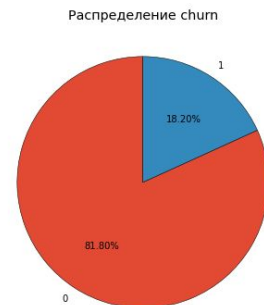
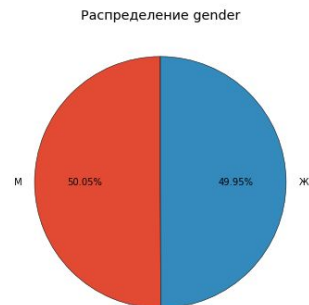
- **Баллы кредитного скоринга:** распределяются от 642 до 1000, среднее значение — 849 баллов, медиана - 853.
- **Возраст:** возраст клиентов от 18 до 86 лет, средний возраст — 42.73 года, медианное значение — 40 лет.
- **Баллы собственности:** от 0 до 9, медиана - 3 балла, среднее значение — 2.63.
- **Баланс:** от 0 до 119 млн. рублей, среднее значение 82 тыс., медианное - 524тысячи.
- **Количество продуктов:** от 0 до 5, , среднее значение — 1.87 продукта, медианное значение — 2.
- **Зарплаты:** указанные значения от 2 546 до 1 395 064 среднее значение — 478 687 , медианное значение — 196 658 рублей.



# Категориальные параметры

- **Города:** Доля клиентов из Ярославля — 58,9%, Ростов — 14,2%, Рыбинск — 26,9%
- **Пол:** мужчин и женщин практически поровну (перекос в 0,1% в пользу мужчин)
- **Активность:** активные клиенты составляют 52.35%, неактивные — 47.65%.
- **Кредитные карты:** 68.04% клиентов имеют кредитную карту, 31.96% — нет.

**! Отток:** среднее значение — **18.2%**, что свидетельствует о достаточно высокой доле ушедших клиентов



# Влияние параметров на отток

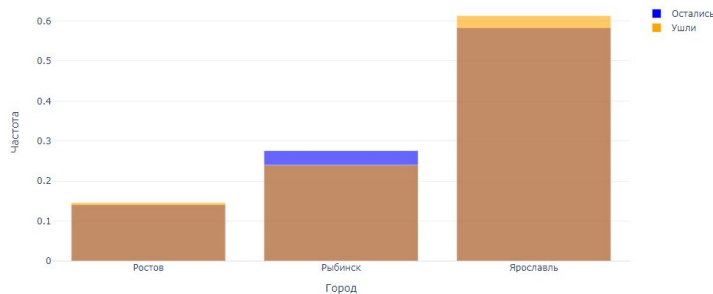
# Определение влияющих признаков

## Город

Средняя доля оттока схожа для Ростова (18.8%) и Ярославля (18.9%), что указывает на примерно одинаковый риск потери клиентов в этих городах.

В Рыбинске наблюдается более низкая доля оттока (16.2%), что может говорить о большей лояльности клиентов в этом городе.

Распределение плотности параметра Город

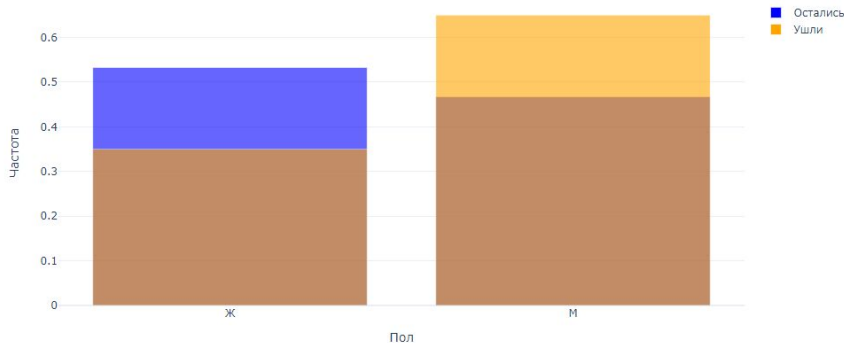


## Пол

У мужчин доля оттока составляет 23.6%, что значительно больше, чем у женщин (12.8%).

Хотя мужчины и женщины представлены равномерно в клиентской базе (50.05% на 49.95%), доля мужчин в общем оттоке составляет 65%, тогда как женщины дают лишь 35%.

Распределение плотности параметра Пол



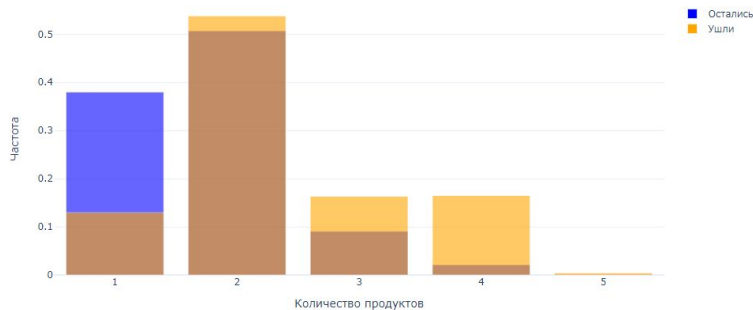


# Определение влияющих признаков

## Продукты

Самые лояльные клиенты - с одним продуктом (7%). У остальных отток выше среднего, с пиком у тех, у кого 4 продукта (отток - 63%), благо, таких клиентов всего 5% от выборки.

Распределение плотности параметра Количество продуктов

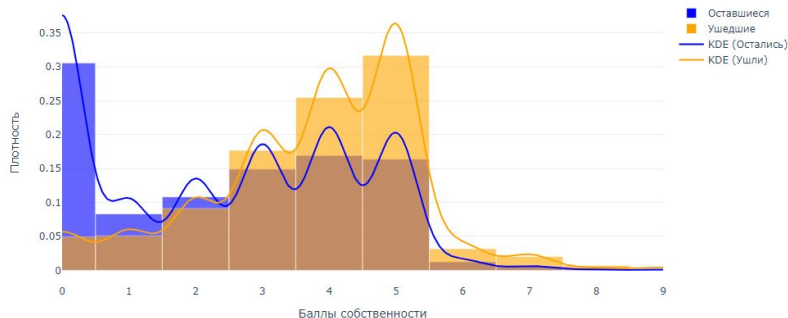


## Баллы собственности

Наибольшая разница плотностей находится на уровне 0 баллов - там наибольшее количество лояльных клиентов.

После 2 баллов отточники начинают превалировать над оставшимися.

Распределение плотности параметра Баллы собственности



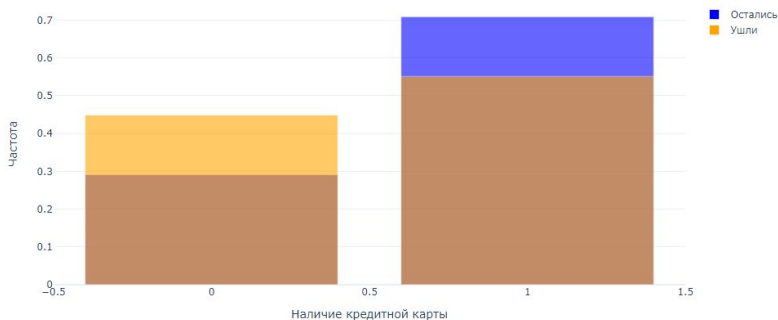
# Определение влияющих признаков

## Наличие кредитной карты

68.04% клиентов имеют кредитную карту, но они дают только 55.1% от общего оттока.

31.96% клиентов не имеют кредитной карты, но они ответственны за 44.9% оттока.

Распределение плотности параметра Наличие кредитной карты

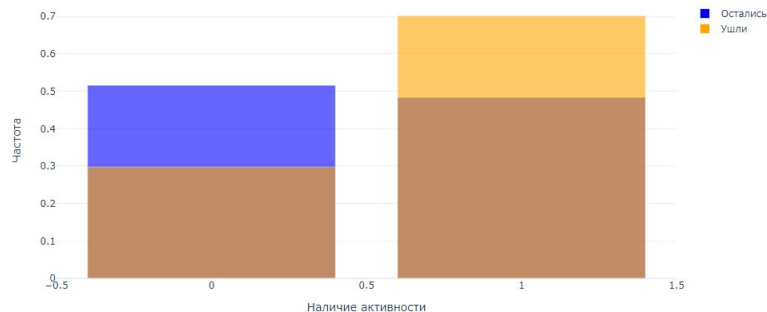


## Наличие активности

Группа активных клиентов составляет 52.35% от общей клиентской базы, но на них приходится 70.2% от общего оттока.

Группа неактивных клиентов занимает 47.65% клиентской базы и ответственна только за 29.8% оттока.

Распределение плотности параметра Наличие активности

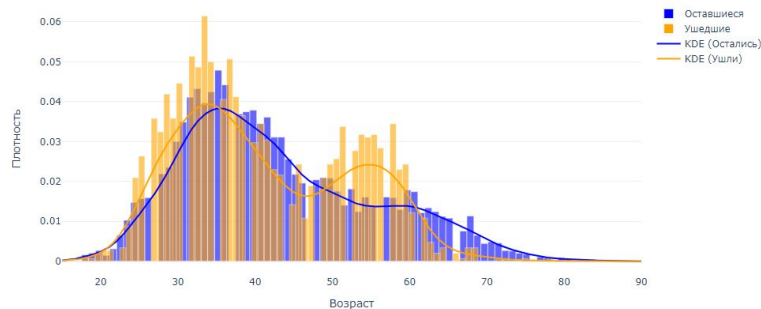


# Определение влияющих признаков

## Возраст

В интервалах 24-37 и 49-59 лет риск ухода выше

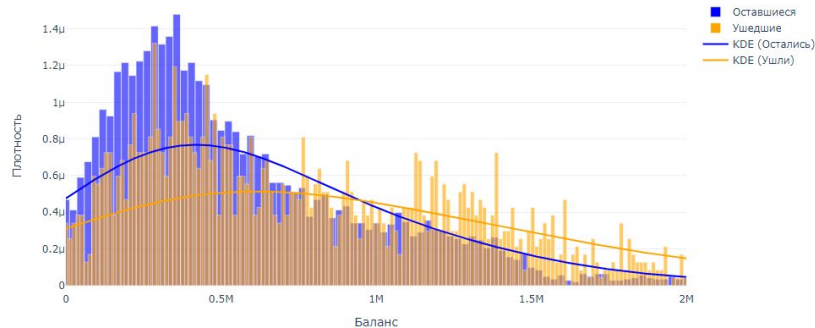
Распределение плотности параметра Возраст



## Баланс

Начиная с баланса примерно в 765 тыс. плотность отточников начинает превышать плотность оставшихся.

Распределение плотности параметра Баланс

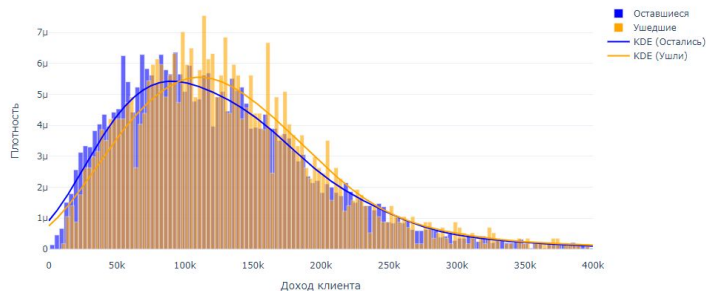


# Определение влияющих признаков

## Доход

Начиная с дохода около 98 тыс. и до 218 тыс плотность вероятности отточников становится незначительно выше плотности остающихся, заем они примерно сравниваются

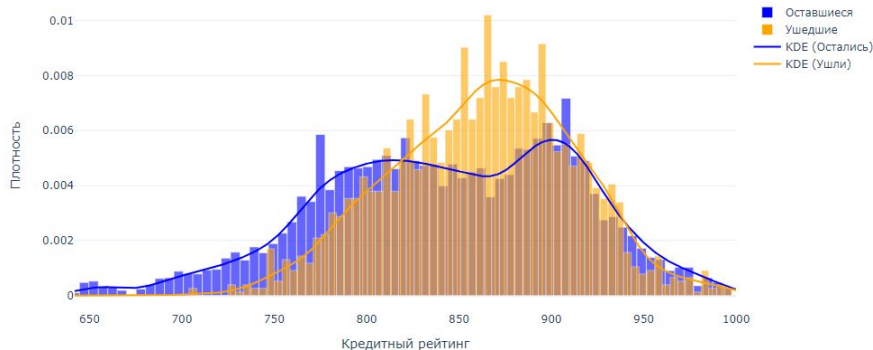
Распределение плотности параметра Доход клиента



## Кредитный рейтинг

Если кредитный рейтинг клиента между 823 и 895 баллов, вероятность ухода возрастает, так как в этом диапазоне оставшихся клиентов меньше, чем тех, кто ушёл.

Распределение плотности параметра Кредитный рейтинг



# Портреты клиентов

# Признаки, характеризующие клиента:

## Склонен к оттоку

- мужской пол
- отсутствие кредитной карты
- наличие активности
- используется более одного продукта
- баллы собственности - 3 и более
- имеет баланс 765 тыс. и выше
- имеет кредитный рейтинг между 823 и 895 баллов

## Не склонен к оттоку

- женский пол
- наличие кредитной карты
- отсутствие активности
- используется всего один продукт
- баллы собственности - 2 и менее баллов
- имеет отсутствующие данные о балансе
- имеет кредитный рейтинг до 810

# Высокоотточные сегменты:

## Сегмент 1

### Параметры сегмента:

- баланс от 765 тысяч
- Мужчина
- Активный клиент

### Характеристики сегмента:

- размер - 702 клиента
- отток в сегменте - 59.69%

## Сегмент 2

### Параметры сегмента:

- Возраст между 24 и 37 годами включительно
- более 2 баллов собственности
- кредитная карта отсутствует

### Характеристики сегмента:

- размер - 750 клиентов
- отток в сегменте - 42.00%

## Сегмент 3

### Параметры сегмента:

- Кредитный рейтинг от 823 до 895
- Возраст между 49 и 59 годами включительно
- 2 и более продуктов

### Характеристики сегмента:

- размер - 524 клиента
- отток в сегменте - 38.93%

# Проверка гипотез



# Проверка гипотез

## Гипотеза 1

**H0 (нулевая гипотеза):** Средний доход клиентов, которые ушли, равен среднему доходу клиентов, которые остались.

**H1 (альтернативная гипотеза):** Средний доход клиентов, которые ушли, отличается от среднего дохода клиентов, которые остались.

**p-значение** теста Манна-Уитни составляет 0.00024, что значительно меньше уровня значимости 0.05.

### Вывод:

Средние доходы между ушедшими клиентами и теми, кто остался, отличаются.

## Гипотеза 2

**H0 (нулевая гипотеза):** Распределение возраста одинаково для ушедших и оставшихся клиентов.

**H1 (альтернативная гипотеза):** Распределение возраста различается между ушедшими и оставшимися клиентами.

Тест Манна-Уитни: U-статистика = 6908213.5. p-value = 0.0000046

### . Вывод:

Возраст влияет на отток.

# Выводы и рекомендации

# Основные выводы

## Общая информация:

- **Количество клиентов в выборке:** 10 тыс.
- **Дубликатов:** явных дубликатов не обнаружено, встречаются повторяющиеся значения userid.
- **Пропущенные значения:** В столбцах **balance** и **age** мы оставили их без изменения, решив, что отсутствие значения - тоже признак.
- **Баллы кредитного скоринга:** распределяются от 642 до 1000, среднее значение — 849 баллов, медиана - 853.
- **Города:** Доля клиентов из Ярославля — 58,9%, Ростова — 14,2%, Рыбинска — 26,9%.
- **Пол:** мужчин и женщин практически поровну (перекос в 0,1% в пользу мужчин)
- **Возраст:** возраст клиентов от 18 до 86 лет, средний возраст — 42.73 года, медианное значение — 40 лет.
- **Баланс:** от 0 до 119 млн. рублей, среднее значение 82 тыс., медианное - 524тысячи.
- **Баллы собственности:** от 0 до 9, медиана - 3 балла, среднее значение — 2.63.
- **Количество продуктов:** от 0 до 5, , среднее значение — 1.87 продукта, медианное значение — 2.
- **Кредитные карты:** 68.04% клиентов имеют кредитную карту, 31.96% — нет.
- **Активность:** активные клиенты составляют 52.35%, неактивные — 47.65%.
- **Зарплаты:** указанные значения от 2 546 до 1 395 064 среднее значение — 478 687 , медианное значение — 196 658 рублей.
- **Отток:** среднее значение — 18.2%, что свидетельствует о достаточно высокой доле ушедших клиентов.

# Основные выводы

## Факторы, влияющие на отток:

1. **Пол:** Мужчины более склонны к оттоку, чем женщины.
2. **Кредитная карта:** Клиенты без кредитной карты (`credit_card = 0`) имеют повышенный риск оттока.
3. **Активность:** Активные клиенты (`last_activity = 1`) чаще уходят из банка.
4. **Количество продуктов:** Клиенты, имеющие 2 продукта и более (`products >= 2`), находятся в группе высокого риска оттока.
5. **Баллы собственности:** Клиенты с баллами собственности более 2 (`equity > 2`) показывают повышенную склонность к оттоку.
6. **Баланс:** Клиенты с балансом от 765 тыс рублей и выше (`balance >= 765000`) имеют высокий риск ухода, клиенты с отсутствующими данными о балансе практически не уходят.
7. **Город:** Клиенты, проживающие в Ярославле и Ростове, более склонны к оттоку, чем клиенты из Рыбинска.
8. **Кредитный рейтинг:** Клиенты с кредитным рейтингом в диапазоне от 823 до 895 (`823 <= score <= 895`) имеют повышенный риск оттока.
9. **Доход клиента:** Клиенты с доходом от 98 до 218 тыс (`98000 <= est_salary <= 218000`) более склонны к оттоку.

# Рекомендации по удержанию

**Сегмент 1** (gender == "M", balance >= 765000, last\_activity == 1)

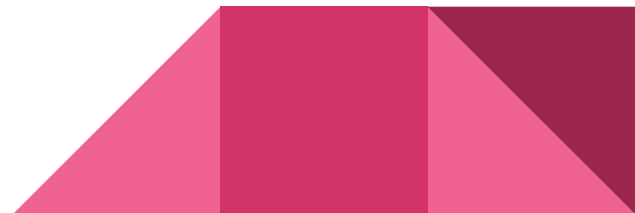
Рекомендация: предложить депозитарные продукты с повышенной ставкой, но с фиксированными длительными сроками хранения

**Сегмент 2** (24 <= age <= 37, equity > 2, credit\_card == 0)

Рекомендация: предложить кредитную карту, предложить льготные программы страхования недвижимости

**Сегмент 3** (823 <= score <= 895, 49 <= age <= 59, products >= 2)

Рекомендация: льготное кредитование, накопительные пенсионных счета



# Дополнительные исследования


## **Обогащение данных:**

- Опросы клиентов для выявления причин оттока.
- Данные о продолжительности сотрудничества.

## **Анализ сезонности:**

- Определить пиковые периоды оттока.

## **Моделирование и тестирование:**

- Предиктивные модели для прогнозирования оттока.
  - A/B тестирование маркетинговых стратегий.
- 

## Итог

Внедрение предложенных рекомендаций позволит банку:

- а. Сократить уровень оттока.
- б. Увеличить лояльность клиентов.
- с. Улучшить финансовые показатели.

**Благодарю за внимание!**

---