

# Выпускной проект

## Банки — Сегментация пользователей по потреблению

Спринт 17

Отчет подготовил  
Лихоузов Кирилл  
8-557-54353  
14/04/2022

# Intro

- **Задача**

Проанализировать клиентов регионального банка и сегментировать пользователей по количеству потребляемых продуктов.

- **Заказчики**

Департамент Розничного бизнеса

- **Цель**

Предложение определенных продуктов для удержания клиентов

## Источник данных

- Источники данных для дашборда: данные о клиентах банка «Метанпром». Банк располагается в Ярославле и областных городах: Ростов Великий и Рыбинск. (таблица **bank\_dataset**)
- База данных, в которой будут храниться агрегированные данные: дополнительные агрегированные таблицы в БД **zen**
- Частота обновления данных: один раз в сутки, в полночь по **UTC**

## Первичный анализ данных

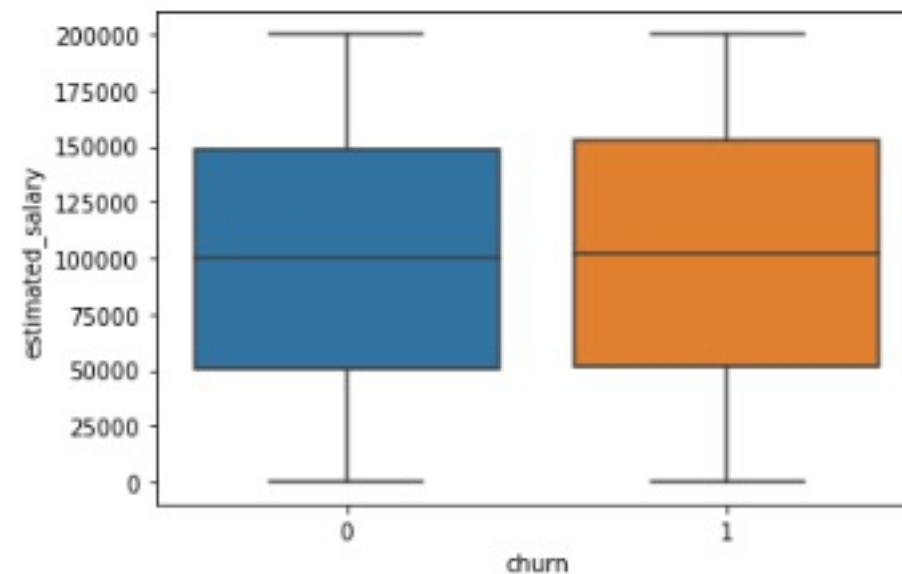
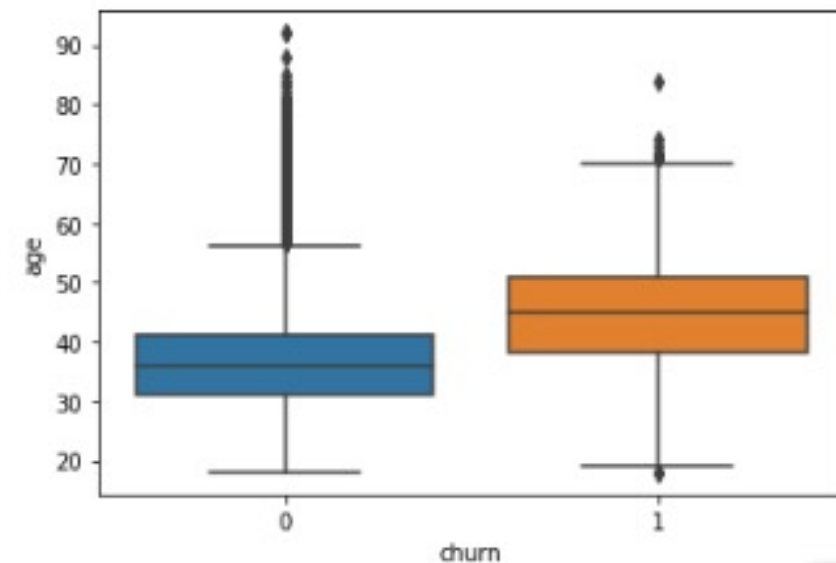
- 10к наблюдений
- имеются пропуски в поле **balance** в Ярославле и Рыбинске. Скорее всего, это связано с тем, что закрытые кредитки не проставляются нулевыми суммами, а фиксируются как пропуски
- отсутствие дублей строк

## Портрет клиента

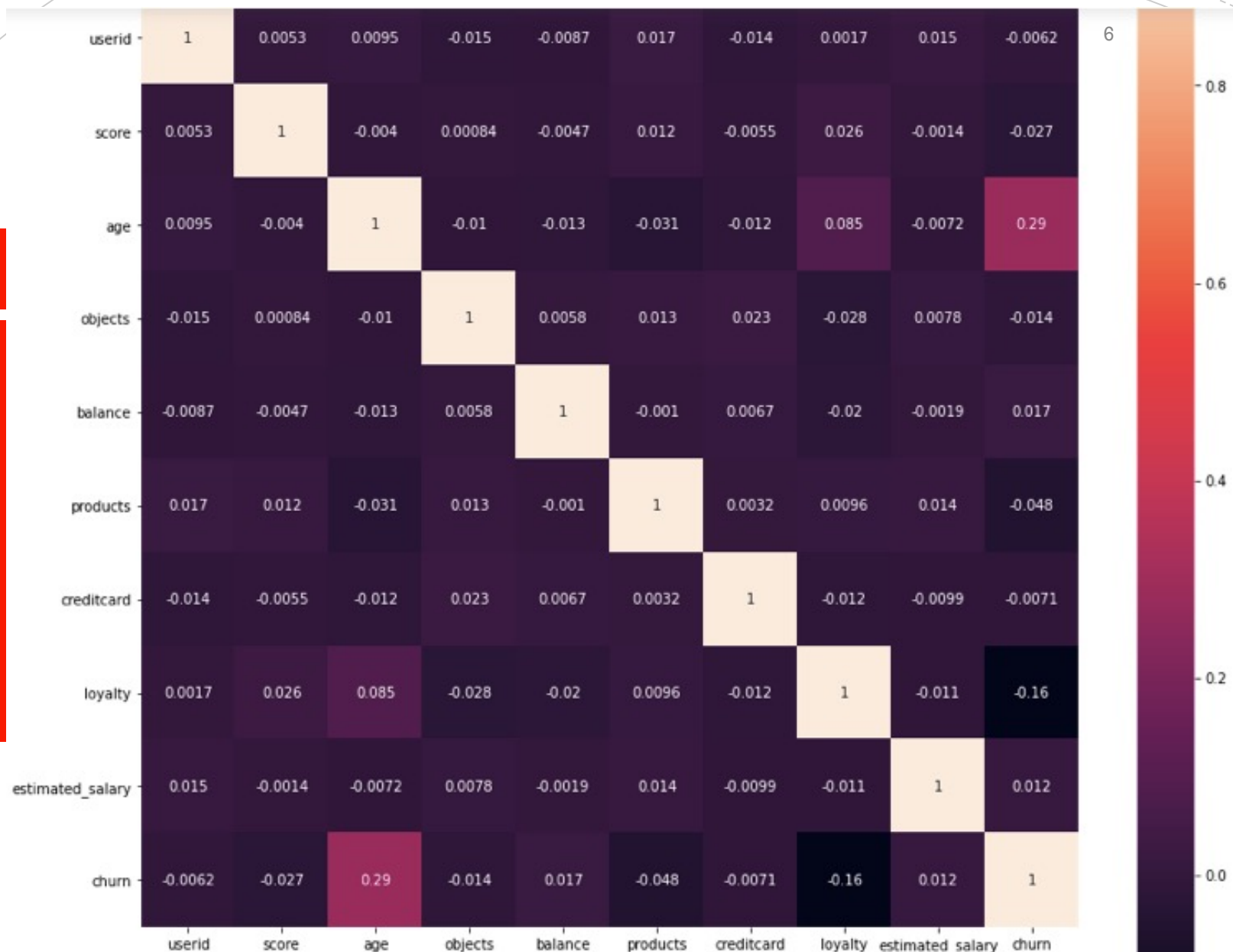
- 38 лет
- сумма на счету - 120к
- имеет 2 продукта
- имеет кредитку

- средний возраст ушедших клиентов 45 лет

- зарплата 100к

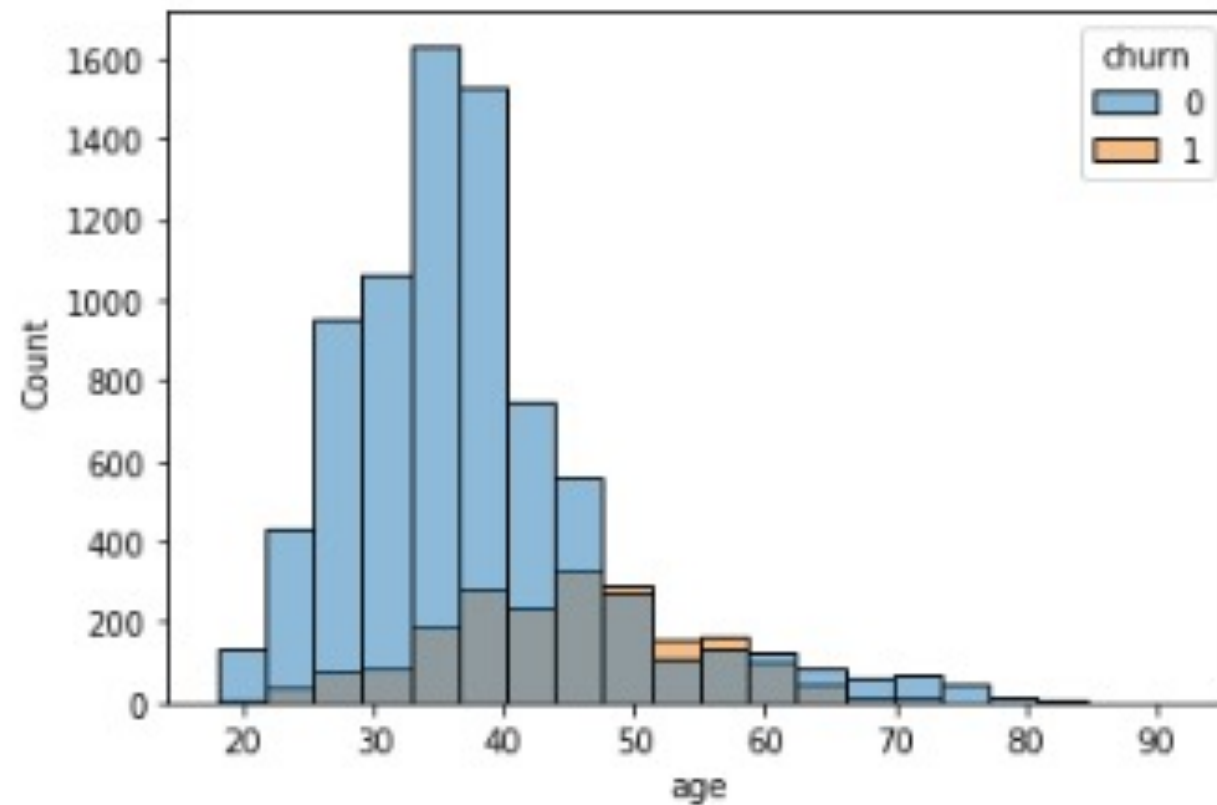


Сильная  
корреляция  
оттока с  
возрастом  
клиента



Возраст ушедших клиентов сильно смещен в большую сторону, относительно тех, кто остался клиентом банка

Молодые  
чаще остаются



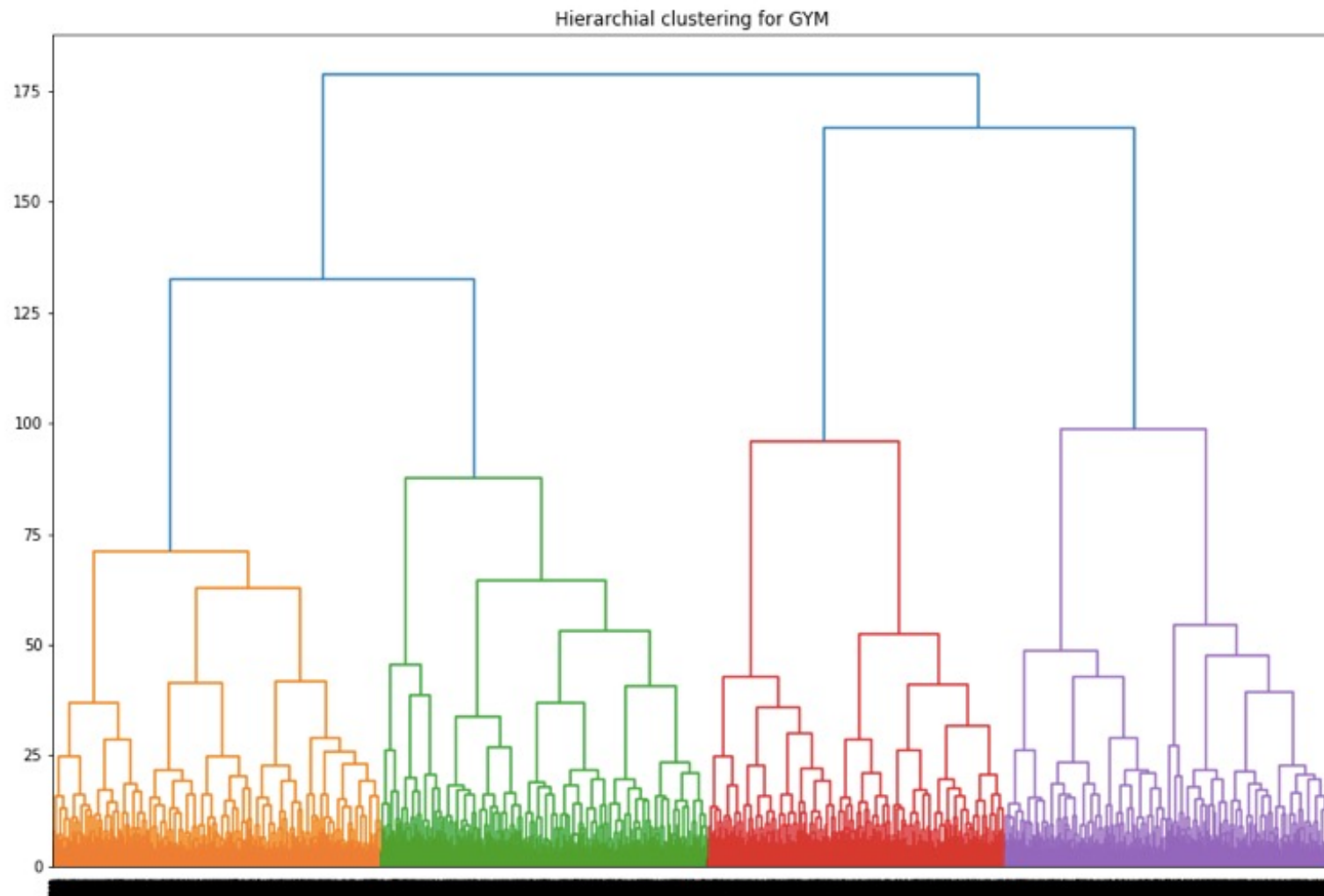
## Проверка гипотез

### *Нулевые гипотезы не были отвержены*

- За нулевые гипотезы ( $H_0$ ) были взяты утверждения:
  - Средние доходы между теми клиентами, которые пользуются двумя продуктами банка, и теми, которые пользуются одним, одинаковые
  - Средние доходы между теми клиентами, у которых есть кредитная карта и теми, у кого ее нет, одинаковые
- Альтернативные гипотезы ( $H_1$ ): Средний доходы различны (соответственно для каждой  $H_0$ )



# Кластеризация: дендрограмма

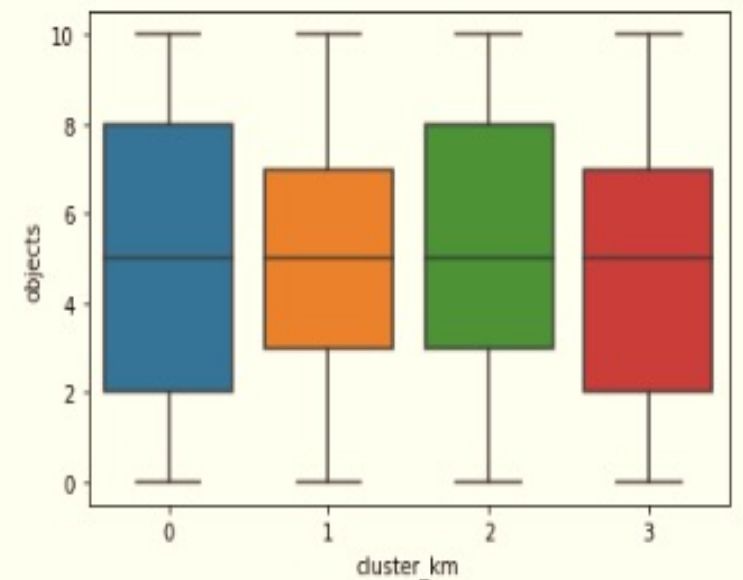
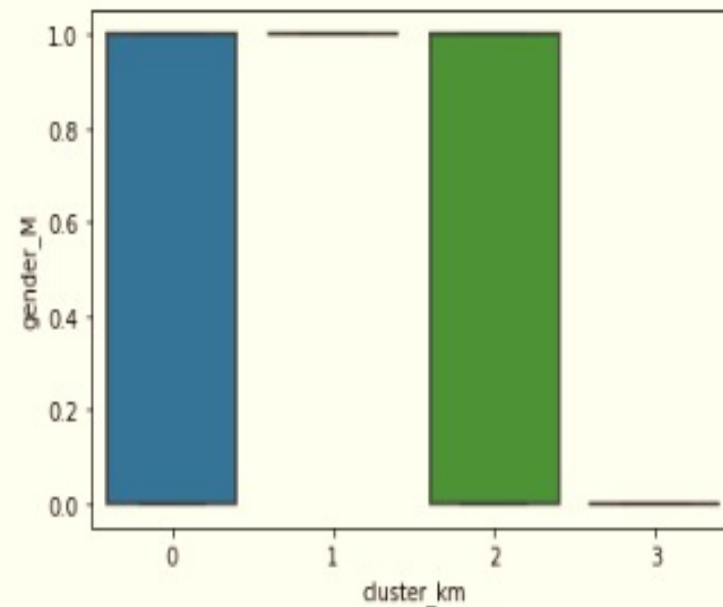
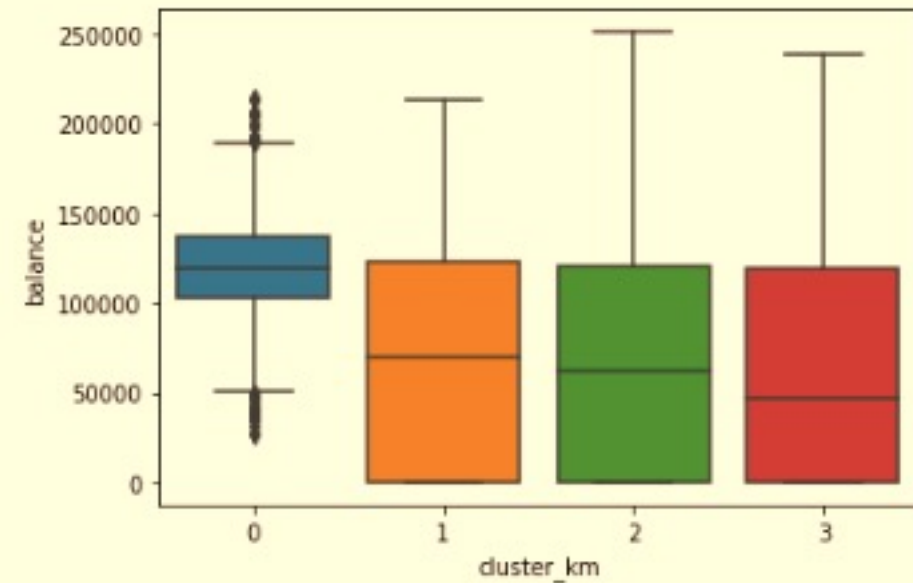


Дендрограмма показывает о достаточности 4 кластеров

## Портрет сегментов

- Сегмент 0:
  - клиенты с балансом 100-140к единиц
  - много клиентов покинувших банк
  - все из Ростова Великого
- Сегмент 1:
  - все мужчины
  - все из Ярославля
- Сегмент 2:
  - все из Рыбинска
- Сегмент 3:
  - все женщины
  - все из Ярославля

# Портрет сегментов



## Summary

- Провели предварительный анализ данных, сформировали портрет клиента
- Проверили две гипотез, обе опровергли
- Провели сегментацию пользователей по потреблению
- основное разбиение по кластерам произошло по признаку города проживания:
  - в кластер 0 попали клиенты из Ростова Великого
  - в кластер 2 - из Рыбинска
  - в кластера 1 и 3 - из Ярославля, причем в кластере 1 - все мужчины, в кластере 3 - все женщины

## Предложения

- Учитывая большой отток клиентов из Ростова Великого, можно им (сегмент 0) предлагать больше программ лояльности, возможно дебетовые карты с повышенным кэшбэком, либо кредитные карты с большим грейс периодом, либо предлагать бонусы по программе Приведи Друга
- У сегмента 3 видим наименьшее медианное значение баланса на счете, при этом зарплата на уровне других сегментов. Поэтому им можно предложить большую закредитованность, то есть программы ипотеки, кредитов, кредитных карт
- Сегменты 2 и 3 без явных аномалий, поэтому по ним предлагается повторить анализ через месяц