

Анализ причин оттока и сегментация клиентов банка "Метанпром"

Выполнила:
Бурдакова Е.Н. 12.07.2022

Содержание

Заказчик исследования: отдел маркетинга Банка

Цель исследования: определение причин оттока клиентов и составление стратегии по сокращению оттока и удержанию клиентов.

Данные для анализа: в нашем распоряжении имеется файл с информацией о клиентах банка в документе имеются 10 000 записей, содержащих следующую информацию:

client_id	name	age	gender	city	balance	products	creditcard	loyalty	estimated_salary	churn
0	Тимошин	42	М	Ростов	1	1	1	1	100000.00	0
1	Ковалев	40	Мужчина	Ростов	1	1	1	1	100000.00	0
2	Тимошин	40	Мужчина	Ростов	1	1	1	1	100000.00	1
3	Тимошин	40	Мужчина	Ростов	1	1	1	1	100000.00	0
4	Тимошин	40	Мужчина	Ростов	1	1	1	1	100000.00	0
5	Тимошин	40	Мужчина	Ростов	1	1	1	1	100000.00	0

• **client_id** — идентификатор клиента;
• **name** — фамилия клиента;
• **age** — возраст;
• **gender** — пол;
• **city** — город;
• **balance** — баланс на счёте;
• **products** — количество продуктов, которыми пользуется клиент;
• **creditcard** — есть ли кредитная карта;
• **loyalty** — лояльный клиент;
• **churn** — ушёл или нет;
• **estimated_salary** — зарплатная планка клиента.

Общие рекомендации отделу маркетинга

- необходимо активизировать направленные на отточных клиентов, в частности:**
 - провести проверку качества обслуживания клиентов и состояния офисов в г. Ростове Великий, чтобы выявить возможные отклонения от стандарта обслуживания в этом городе;
 - совершать ред. действия (предложить), которые были бы интересны клиентам среднего возраста, например типичные вещи: полезные им:
 - сидки в косметическом салоне;
 - сидки на выходные;
 - сидки на определённые виды товаров в продуктовых магазинах.

- уделить внимание клиентам с балансом выше 100 тысяч, например дополнительные бонусы в виде процента на остаток при подержании заданного баланса.

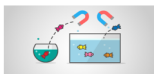
- увеличить** число клиентов, у которых будут соблюдаться маркеры характеризующие слабый отток:

- добавить в приложение банка опции, заставляющие клиента регулярно заходить в приложение и совершать какие-либо действия:
 - бонусные игры, со сбором фишек/звёзд, которые можно в дальнейшем обменивать на сидки или иные льготные предложения второго продукта.
- клиентам пользующимся одним продуктом банка постараться продать ещё один продукт, т.к. клиенты с двумя продуктами менее склонны к оттоку, предложив по второму продукту должны учитывать потребности клиентов

- уделить внимание контролю качества, удовлетворённости от использования продуктов**

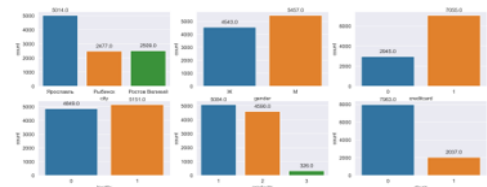
- разработка сценария для обзора, собирающего информацию о удовлетворённости от использования и пожеланий клиентов, это позволит оптимизировать действующую и разработать новые продукты.
- организовать сбор обратной связи удовлетворённости клиентов из зоны риска.

Реализация всех вышеописанных рекомендаций потребует тщательной проработки профильными подразделениями банка



Общий «вид» клиентов банка

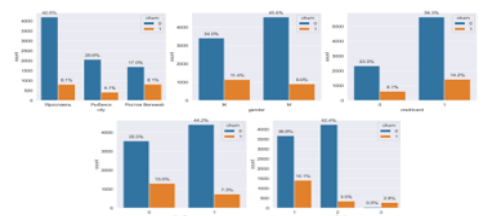
Количество клиентов в разрезе категориальных признаков



- в Ростове и в Ростове Великом примерно поровну клиентов, в Ростове продукт больше;
- мужчин noticeably больше, чем женщин;
- младше 40 лет значительно больше клиентов не пользующихся картой;
- клиенты совершающие какие-либо действия в приложении и не получающие этого критически важного;
- в выборке есть клиенты с 1, 2, 3 и 4 продуктами банка, преобладают с 1 или 2;
- отток довольно большой, около 20%.

Портрет клиента склонного к оттоку

Количество клиентов в разрезе категориальных признаков и оттока

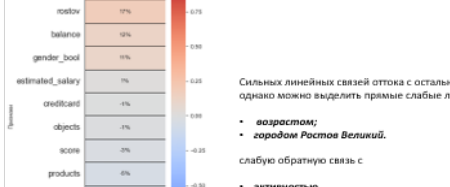


Критерии клиента в «зоне риска»:

- самая высокая доля оттока в Ростове Великом;
- доля оттока среди тех, кто совершает действия в приложении заметно ниже, чем среди непользующихся клиентами;
- несмотря на то, что клиентов с 3-4 продуктами мало, доля оттока у них выше, чем у держателей;
- клиенты с балансом выше 100 тысяч;
- высокая доля оттока у клиентов с одним продуктом;

Корреляционный анализ

Тепловая карта по признакам



Сильных линейных связей оттока с остальными параметрами нет, однако можно выделить прямые слабые линейные связи с:

- возрастом;
- городом Ростов Великий;
- слабую обратную связь с активностью

Матрица корреляций phi_k

Можно отметить следующие корреляции по оттоку:

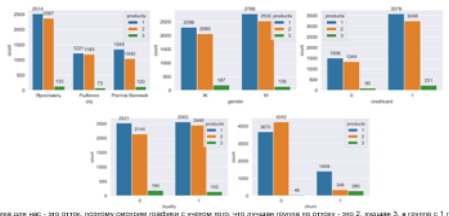
- с количеством продуктов 57%;
- возрастом 43%.

В несколько меньшей степени с:

- городом Ростов Великий 27%;
- активностью 24%.

Анализ клиентов в разрезе количества продуктов

Количество клиентов в разрезе категориальных признаков и продуктов



Конечная метрика для нас - это отток, поэтому смотрим графики с учётом того, что лучшей группой по оттоку - это 2, худшей 3, а группа с 1 продуктом значимо оттока также оставшимся лучше.

- в группе 2 чуть чаще пользуются приложением;
- меньше всего клиентов с двумя продуктами проживают в Ростове Великом;
- 1-2 продуктами чаще пользуются женщины, 3-4 продуктами - мужчины;
- во всех группах пользующихся кредитной картой больше, чем клиентов не имеющих данного продукта.

* группы с 3 и 4 продуктами объединены в одну

Доли клиентов с потенциально опасным значением данных признаков



name	size	avg_churn
0 женщины	4543	0.25
1 старые 42	2894	0.42
2 из Ростова	2509	0.32
3 высокий баланс	4799	0.25
4 не 2 продукта	5410	0.31

Во всех группах довольно высокий отток, выше общего по данным (20%). Схожи по доле оттока сегменты: Женщины и высокий баланс, Ростов Великий и не 2 продукта.

Рекомендации в разрезе сегментов

- клиентам из Ростова Великого с одним продуктом предложить оформить второй продукт, продуктом должно быть что-то популярное и соответствующее их потребностям (выявленным менеджером) скорее всего вклад или КК (если нет)
*предварительно убедившись, что их устраивает качество обслуживания и отработав возможные возражения;
- клиенты старше 42 лет скорее всего имеют детей-подростков, автомобиль, дом/квартиру под все это подходит ряд программ страхования, предложение которых будет релевантно, также есть вероятность, что в этом возрасте клиенты уже обзаводятся сбережениями и вклад на льготных условиях будет популярен;
- женщинам с балансом выше 100 000 целесообразно дополнительно оформить вклад с более привлекательными условиями

Заключение

Зона риска:

- мужчины;
- клиенты в возрасте 44-45 лет и выше;
- лица проживающие в Ростове Великом;
- клиенты с балансом выше 100 тысяч;
- не используют приложение;
- пользователи 1 или 3-4 продуктов.

Надёжный клиент:

- мужчины;
- клиенты в возрасте до 44 лет;
- клиенты регулярно совершающие действия в приложении банка;
- пользователи двух продуктов банка.

Ознакомиться с рекомендациями заказчика



Спасибо за внимание!

Заказчик исследования: отдел маркетинга Банка

Цель исследования: определение причин оттока клиентов и составление стратегии по сокращению оттока и удержанию клиентов.

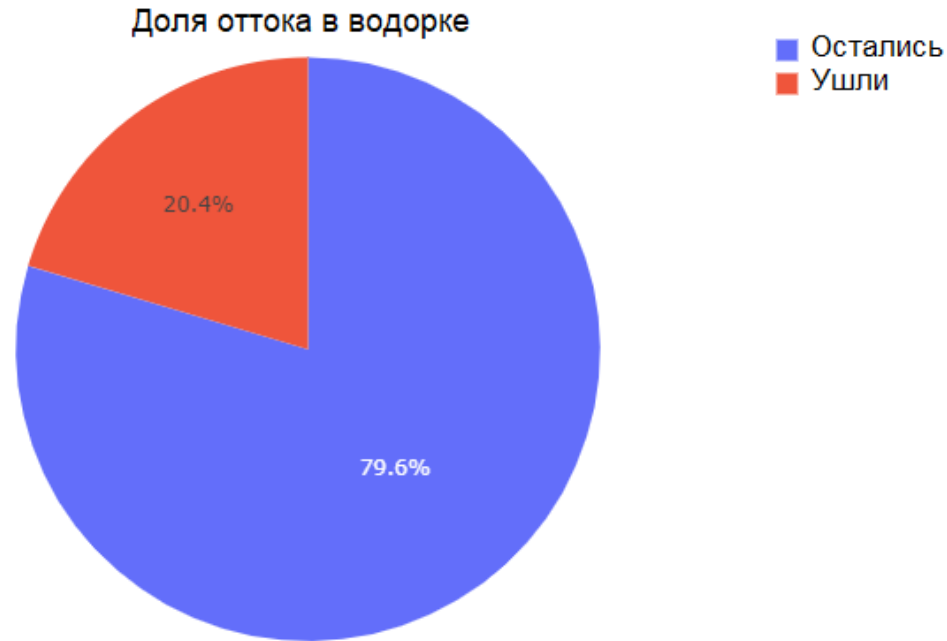
Данные для анализа: в нашем распоряжении имеется файл с информацией о клиентах банка в документе имеются 10 000 записей, содержащих следующую информацию:

	userid	score	City	Gender	Age	Objects	Balance	Products	CreditCard	Loyalty	estimated_salary	Churn
0	15677338	619	Ярославль	Ж	42	2	NaN	1	1	1	101348.88	1
1	15690047	608	Рыбинск	Ж	41	1	83807.86	1	0	1	112542.58	0
2	15662040	502	Ярославль	Ж	42	8	159660.80	3	1	0	113931.57	1
3	15744090	699	Ярославль	Ж	39	1	NaN	2	0	0	93826.63	0
4	15780624	850	Рыбинск	Ж	43	2	125510.82	1	1	1	79084.10	0

- **userid** — идентификатор пользователя,
- **score** — баллы кредитного скоринга,
- **City** — город,
- **Gender** — пол,
- **Age** — возраст,
- **Objects** — количество объектов в собственности,
- **Balance** — баланс на счёте,
- **Products** — количество продуктов, которыми пользуется клиент,
- **CreditCard** — есть ли кредитная карта,
- **Loyalty** — активный клиент,
- **Churn** — ушёл или нет.
- **estimated_salary** — заработная плата клиента,

Доля оттока клиентов банка превышает 20%.
Столь высокое значение данного показателя вызывает серьезное беспокойство.

В данной работе мы проанализируем причины оттока и
сформулируем рекомендации по его сокращению

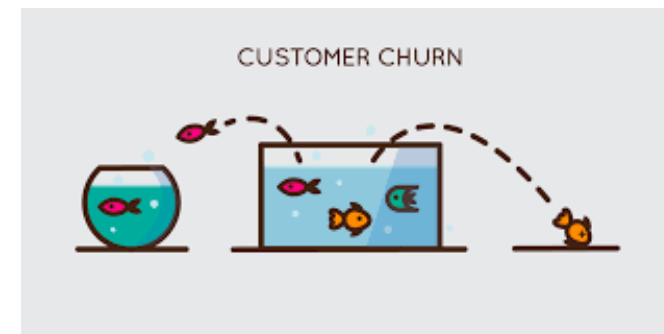


Далее представлены результаты анализа, проведенные на основании имеющихся данных

Общие рекомендации отделу маркетинга

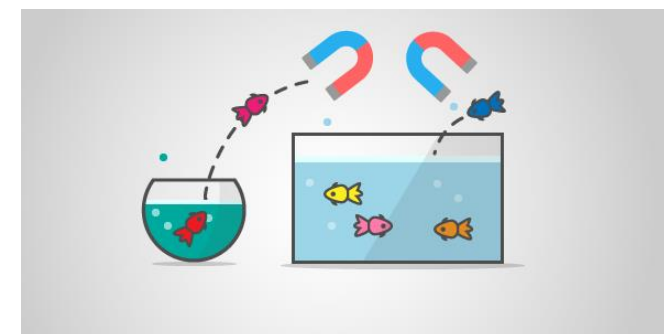
1 - необходимы активности направленные на отточных клиентов, в частности:

- провести проверку качества обслуживания клиентов и состояния офисов в г. Ростов Великий, чтобы выявить возможные отклонения от стандартов обслуживания в этом городе;
- сформировать ряд акционных предложений, которые были бы интересны женщинам среднего возраста, например типичные вещи полезные им:
 - скидки в косметических салонах;
 - скидки магазинах одежды;
 - скидки на определенные виды товаров в продуктовых магазинах.
- уделить внимание клиентам с балансом выше 100 тысяч, например дополнительные бонусы в виде процента на остаток при поддержании заданного баланса.



2 - увеличивать число клиентов, у которых будут соблюдаться маркеры характеризующие слабый отток:

- добавить в приложение банка опции, заставляющие клиента регулярно заходить в приложение и совершать какие - либо действия.
 - бонусные игры, со сбором фишек/рыбок, которые можно в дальнейшем обменивать на скидки и/или льготное оформление второго продукта.
- клиентам пользующимся одним продуктом банка постараться продать еще один продукт, т.к. клиенты с двумя продуктами менее склонны к оттоку, предложения по второму продукту должны учитывать потребности клиентов



3 - уделить внимание контролю качества, удовлетворенности от использования продуктов

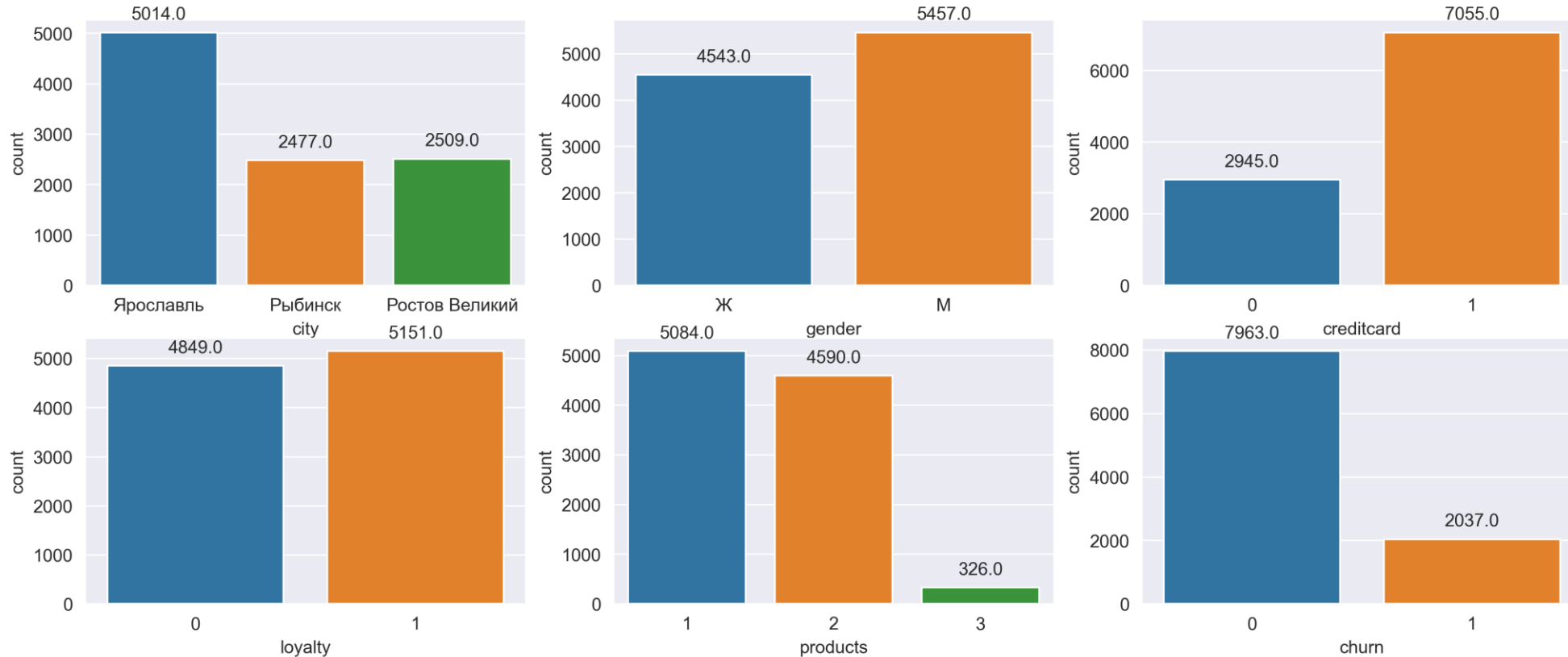
- разработка скрипта для обзвона, собирающего информацию о удовлетворенности от использования и пожеланий клиентов, это позволит оптимизировать действующие и разработать новые продукты.
- организовать сбор обратной связи удовлетворенности клиентов из зоны риска.



Реализация всех вышеописанных рекомендаций требует тщательной проработки профильными подразделениями банка

Общий «вид» клиентов банка

Количество клиентов в разрезе категориальных переменных

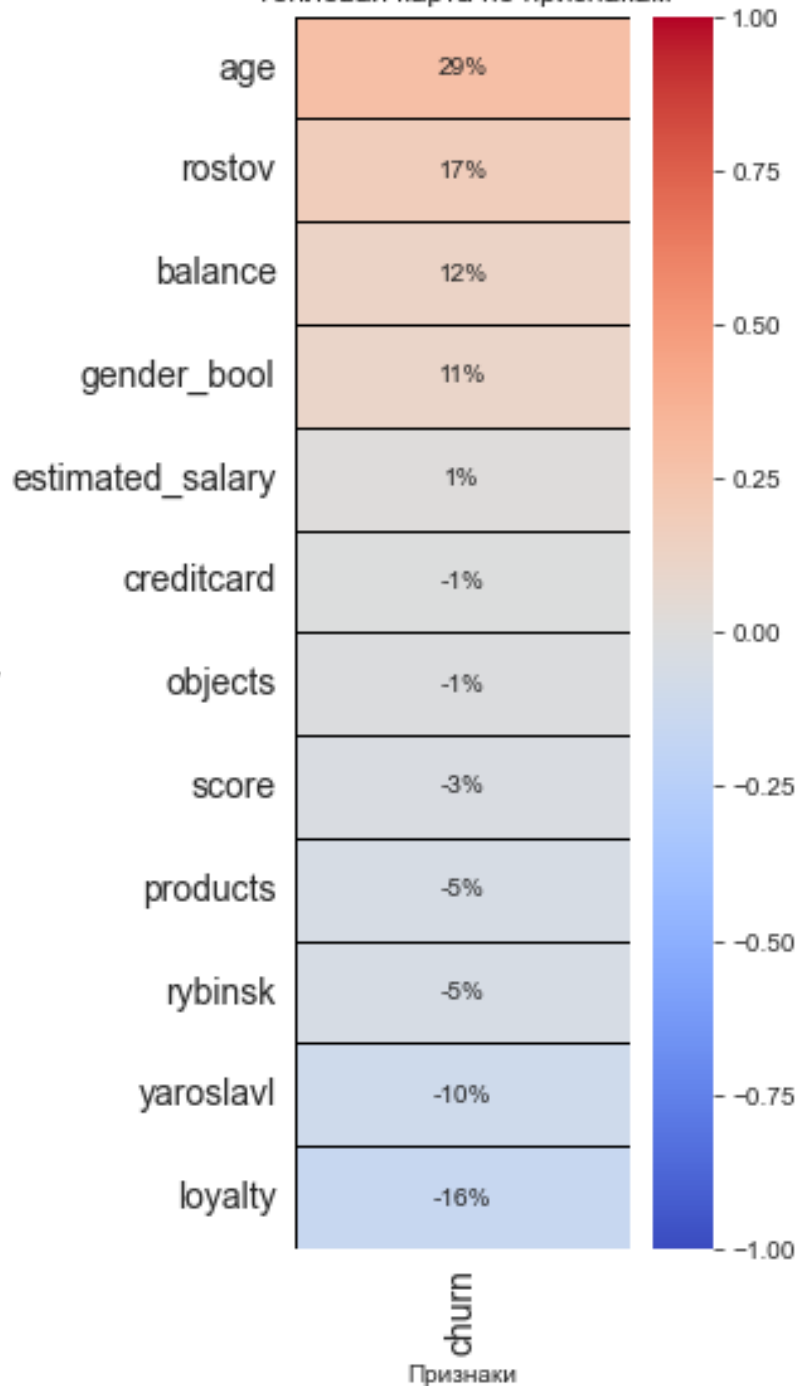


- в Рыбинске и Ростове Великом примерно поровну клиентов, в Ярославле гораздо больше;
- мужчин немного больше чем женщин;
- владельцев КК значительно больше чем клиентов не пользующихся картой;
- клиентов совершающих какие-либо действия в приложении и не делающих этого практически поровну;
- в выборке есть клиенты с 1,2,3 и 4 продуктами банка, преобладают с 1 или 2;
- **отток довольно большой, около 20%.**



- большинство клиентов имеют высокий кредитный рейтинг в среднем около 650;
- возраст клиентов банка чаще всего от 32 до 44 лет, но встречаются граждане 18 и 92 лет;
- распределение кредитного рейтинга, возраста и баланса походит на нормальное, хотя имеются выбросы;
- распределение рассчитанного дохода скорее всего равномерное.

Тепловая карта по признакам



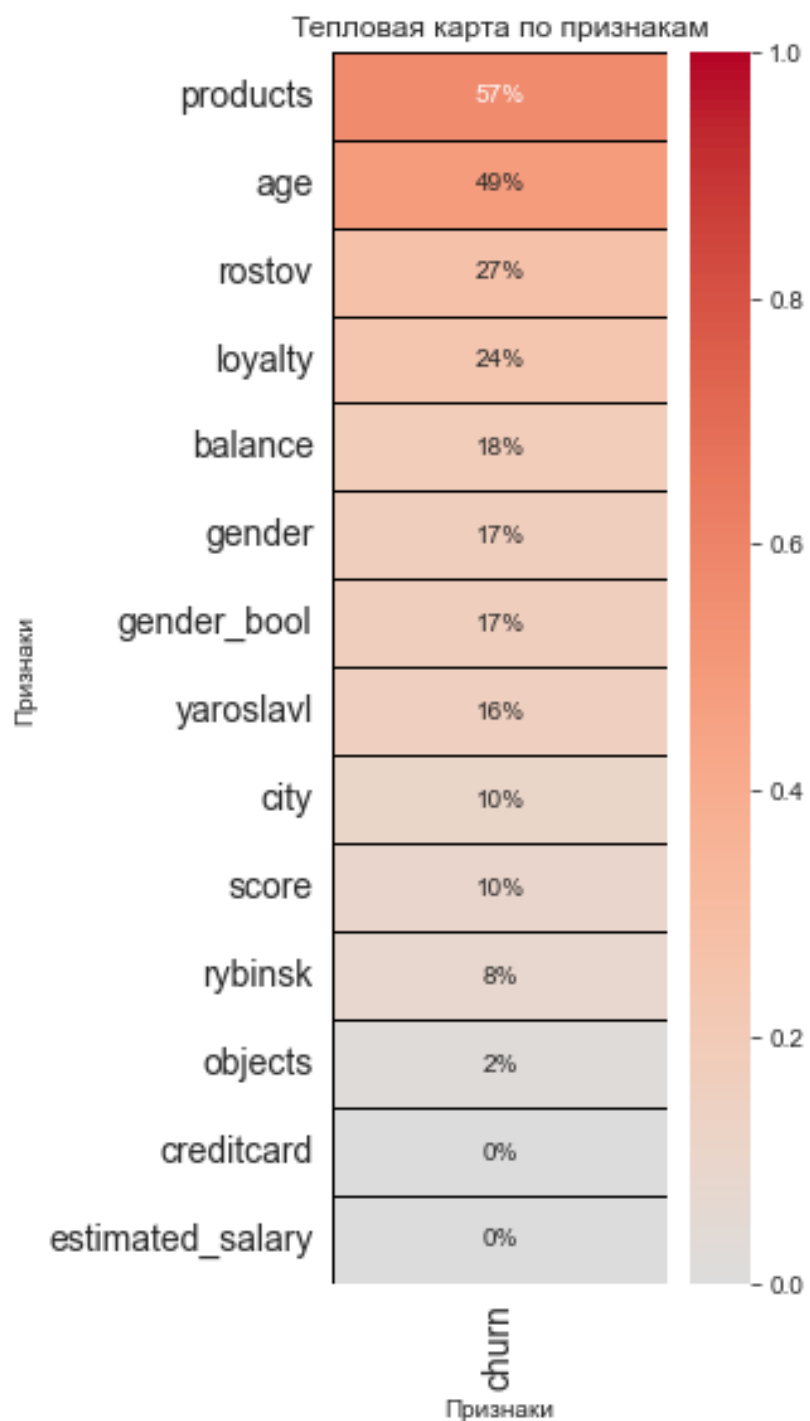
Корреляционный анализ

Сильных линейных связей оттока с остальными параметрами нет, однако можно выделить прямые слабые линейные связи с:

- **возрастом;**
- **городом Ростов Великий.**

слабую обратную связь с

- **активностью**



Матрица корреляций ρ_{ik}

Можно отметить следующие корреляции по оттоку:

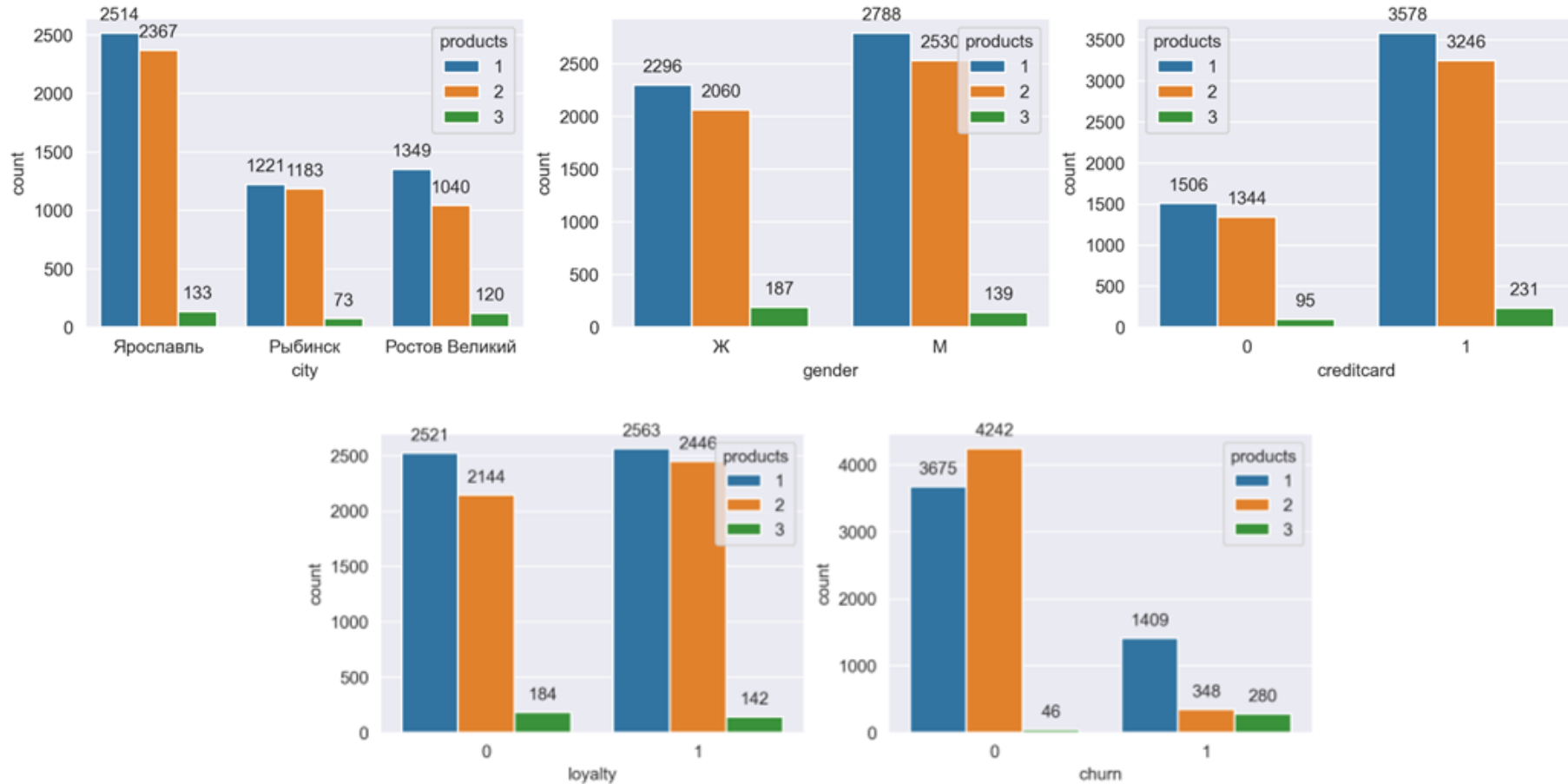
- с количеством продуктов **57%**;
- возрастом **49%**.

В несколько меньшей степени с:

- городом Ростов Великий **27%**;
- активностью **24%**.

Анализ клиентов в разрезе количества продуктов

Количество клиентов в разрезе категориальных переменных и продуктов

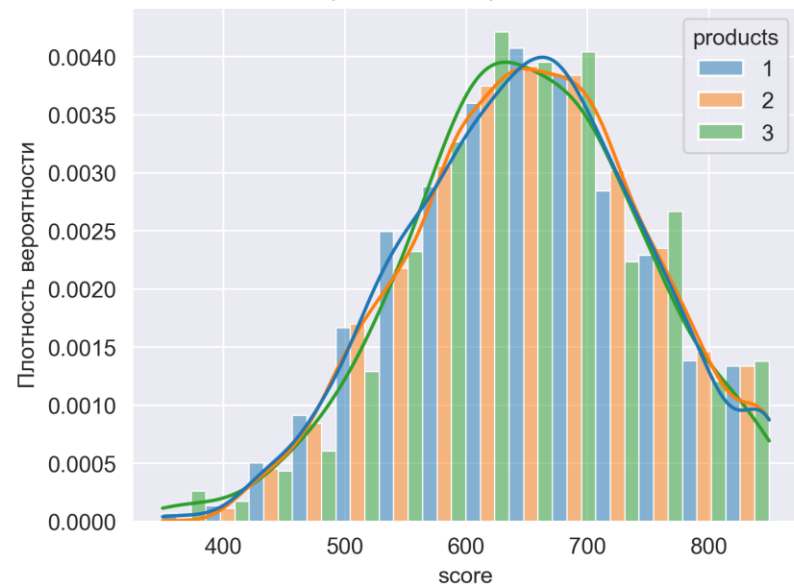


Ключевая метрика для нас - это отток, поэтому смотрим графики с учетом того, что лучшая группа по оттоку - это 2, худшая 3, в группе с 1 продуктом значение оттока также оставляет желать лучшего.

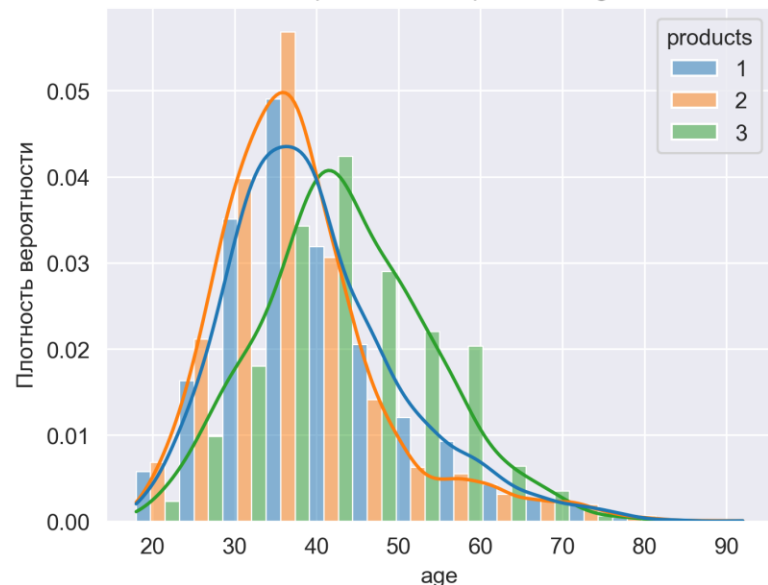
- в группе 2 чуть чаще пользуются приложением;
- меньше всего клиентов с двумя продуктами проживают в Ростове Великом;
- 1-2 продуктами чаще пользуются мужчины, 3 - 4 женщины;
- во всех группах пользователей кредитных карт больше чем клиентов не имеющих данного продукта.

* группы клиентов с 3 и 4 продуктами объединены в одну

Распределение признака score



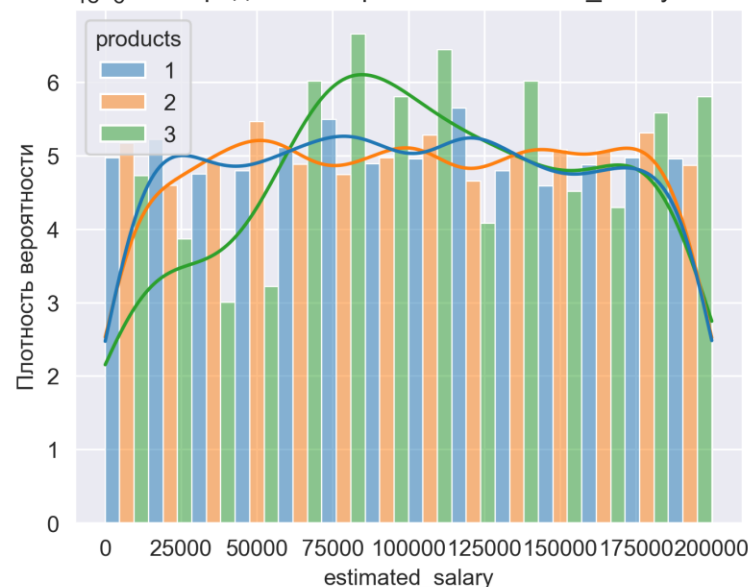
Распределение признака age



Распределение признака balance



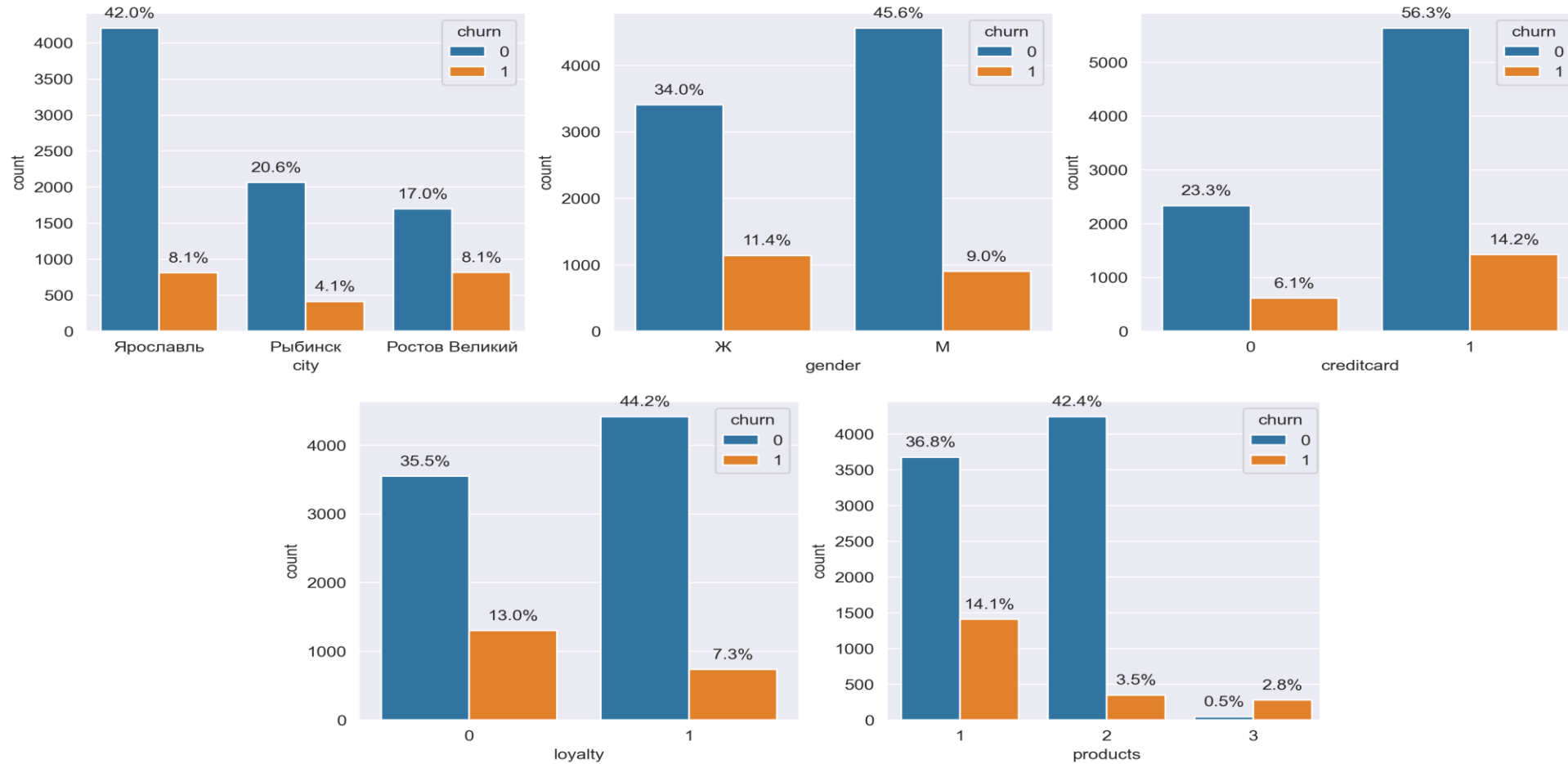
Распределение признака estimated_salary



- в группе с 3-4 продуктами люди более старшего возраста;
- клиенты с двумя продуктами и минимальным чаще попадают в возрастной сегмент 30-40 лет;
- распределение признака дохода равномерное, остальных признаков близко к нормальному.

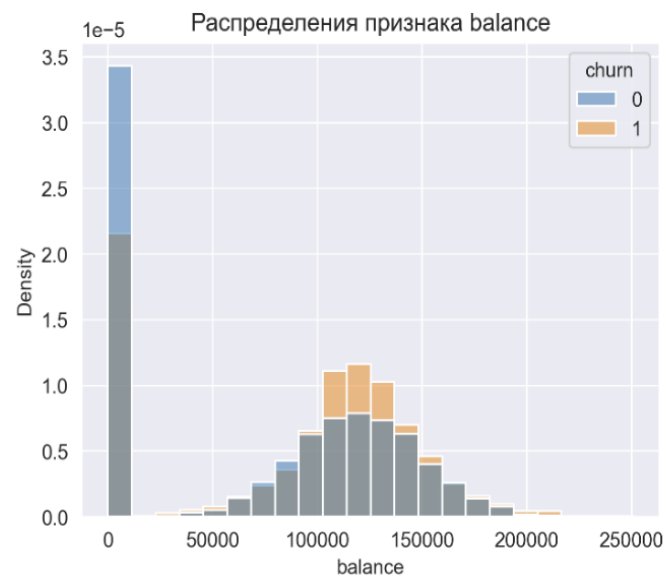
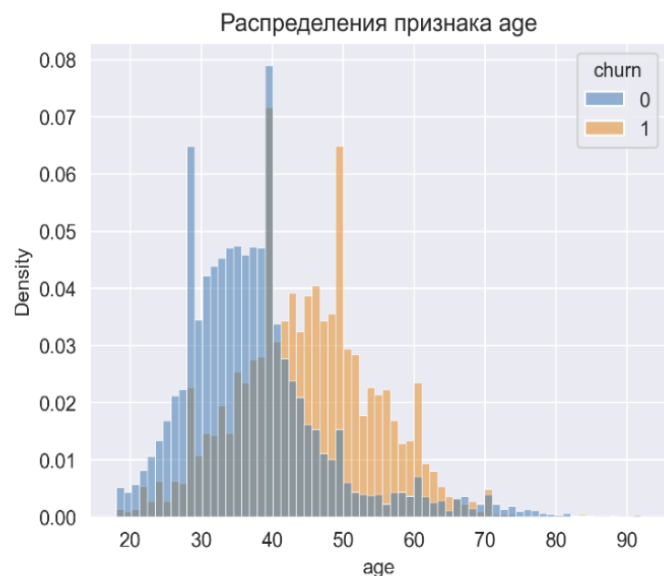
Портрет клиента склонного к оттоку

Количество клиентов в разрезе категориальных переменных и оттока



Критерии клиента в «зоне риска»:

- самая высокая доля оттока в Ростове Великом;
- доля оттока среди тех кто совершает действия в приложении заметно ниже чем среди малоактивных клиентов;
- несмотря на то, что клиентов с 3-4 продуктами мало, доля оттока в них выше доли удержания;
- женщины чуть более склонны к оттоку;
- высокая доля оттока у клиентов с одним продуктом;



Критерии клиента в «зоне риска»:

- возраст примерно 42 и выше;
- остаток на балансе выше 100 тысяч.

Статистический анализ. Построение модели прогнозирования оттока клиентов

В ходе работы были сформулированы и проверены следующие гипотезы:

$$\begin{cases} H_0 : \text{Средний доход отточных и неотточных клиентов равен} \\ H_1 : \text{Средний доход отточных и неотточных клиентов различается} \end{cases}$$

$\alpha=5\%$ критический уровень статистической значимости.

$$\begin{cases} H_0 : \text{Средний возраст отточных и неотточных клиентов равен} \\ H_1 : \text{Средний возраст отточных и неотточных клиентов различается} \end{cases}$$

$\alpha=5\%$ критический уровень статистической значимости.

Для проверки гипотез в обоих случаях применялся ttest Стьюдента. Результаты получились следующими:

- средний доход отточных и неотточных клиентов равен;
- средний возраст отточных и неотточных клиентов различается

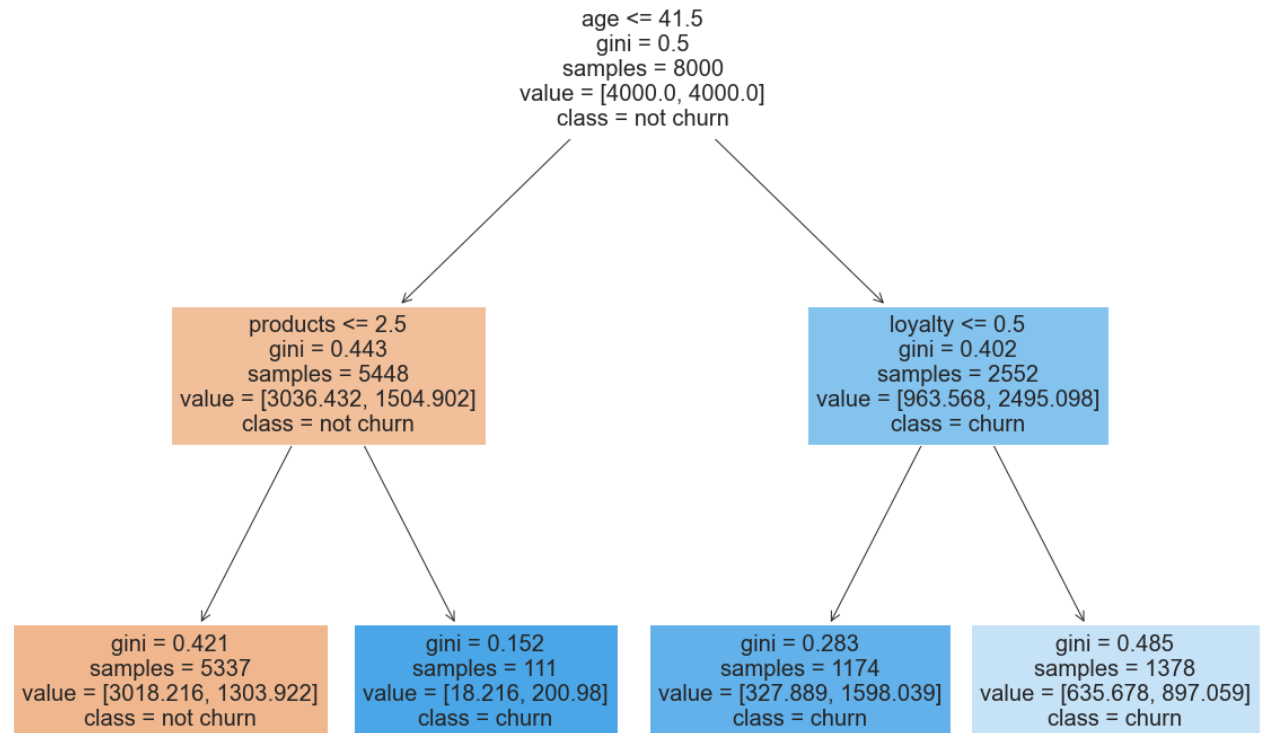
Построены модели прогнозирования оттока клиентов:

- логистическая регрессия;
- случайный лес;
- дерево принятия решений.

Лучший результат показала модель “Случайный лес”
accuracy = 0.87 precision = 0.75 recall = 0.52 F1 = 0.62

* Чем ближе значение метрик к 1, тем точнее работает модель, ключевая метрика F1

Случайный лес - это сводный алгоритм состоит из множества деревьев.
Здесь приведен пример иллюстрации дерева принятия решений.



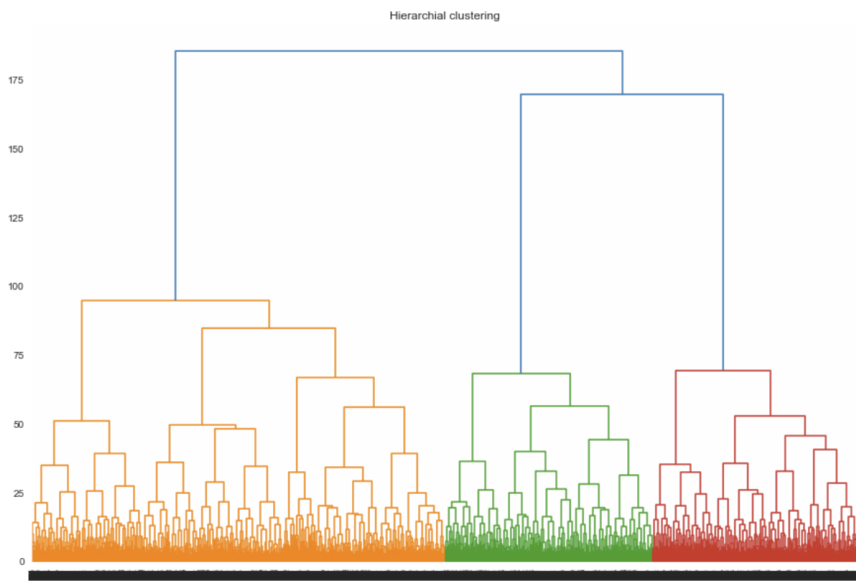
Кластеризация

Для кластеризации спользовался метод `KMeans`,

С помощью дендограммы выделено три кластера.

Качество проведенной кластеризации не высокое `Silhouette_score` = 0,2,

Средние значения характеристик по кластерам не противоречат и подтверждают сделанные ранее выводы о портретах клиентов.



Средние значения параметров в разрезе по оттоку и по итогам кластеризации

	churn	0	1
score		651.85	645.35
age		37.41	44.84
objects		5.03	4.93
balance		72745.30	91108.54
products		1.54	1.45
creditcard		0.71	0.70
loyalty		0.55	0.36
estimated_salary		99738.39	101465.68
yaroslavl		0.53	0.40
rybinsk		0.26	0.20
rostov		0.21	0.40
gender_bool		0.43	0.56

	cluster_km	0	1	2
score		649.67	651.45	651.33
age		38.51	39.77	38.89
objects		5.00	5.01	5.03
balance		62092.64	119730.12	61818.15
products		1.53	1.51	1.54
creditcard		0.71	0.71	0.69
loyalty		0.52	0.50	0.53
estimated_salary		99899.18	101113.44	99440.57
churn		0.16	0.32	0.17
yaroslavl		1.00	0.00	0.00
rybinsk		0.00	0.00	1.00
rostov		0.00	1.00	0.00
gender_bool		0.45	0.48	0.44

Доли клиентов с потенциально опасным значением данных признаков



	name	size	avg_churn
0	женщины	4543	0.25
1	старше 42	2894	0.42
2	из Ростова	2509	0.32
3	высокий баланс	4799	0.25
4	не 2 продукта	5410	0.31

Во всех группах довольно высокий отток, выше общего по данным (20%). Схожи по доле оттока сегменты: Женщины и высокий баланс, Ростов Великий и не 2 продукта.

Сегменты для первоочередного воздействия



Имеем три сегмента с крайне высоким оттоком

name	size	avg_churn
Ростов не 2 продукта	1469	0.47
старше 42	2894	0.42
женщины с высоким балансом	2148	0.30

Необходимы целевые воздействия на каждую из этих групп.
В первую очередь я бы рекомендовала направить работу на группы «старше 42» или «Ростов не 2 продукта».

Сокращение оттока хотя бы в одной из групп значительно сократит общий отток и спасет Банк.

Далее даны рекомендации по каждому сегменту

Рекомендации в разрезе сегментов

- клиентам из Ростова Великого с одним продуктом предложить оформить второй продукт, продуктом должно быть что - то популярное и соответствующее их потребностям (выявленными менеджером) скорее всего вклад или КК (если нет)
 - *предварительно убедившись, что их устраивает качество обслуживания и отработав возможные возражения;
- клиенты старше 42 лет скорее всего имеют детей-подростков, автомобиль, дом/квартиру под все это подходит ряд программ страхования, предложение которых будет релевантно, также есть вероятность, что в этом возрасте клиенты уже обзаводятся сбережениями и вклад на льготных условиях будет полезен;
- женщинам с балансом выше 100 000 целесообразно дополнительно оформить вклад с более привлекательными условиями

Заключение

Зона риска:

- женщины;
- клиенты в возрасте 44-45 лет и выше;
- лица проживающие в Ростове Великом;
- клиенты с балансом выше 100 тысяч;
- не используют приложение;
- пользователи 1 или 3-4 продуктов.

Надежный клиент:

- мужчины;
- клиенты в возрасте до 44 лет;
- клиенты регулярно совершающие активности в приложении банка;
- пользователи двух продуктов банка.

Ознакомиться с рекомендациями заказчику

Общие рекомендации отделу маркетинга

1 - необходимы активности направленные на отточных клиентов, в частности:

- провести проверку качества обслуживания клиентов и состояния офисов в г. Ростов Великий, чтобы выявить возможные отклонения от стандартов обслуживания в этом городе;
- сформировать ряд акционных предложений, которые были бы интересны женщинам среднего возраста, например типичные вещи полезные им:
 - скидки в косметических салонах;
 - скидки магазинах одежды;
 - скидки на определенные виды товаров в продуктовых магазинах.
- уделить внимание клиентам с балансом выше 100 тысяч, например дополнительные бонусы в виде процента на остаток при поддержании заданного баланса.

2 - увеличивать число клиентов, у которых будут соблюдаться маркеры характеризующие слабый отток:


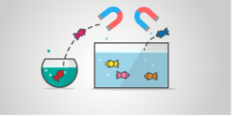
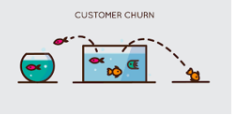
- добавить в приложение банка опции, заставляющие клиента регулярно заходить в приложение и совершать какие-либо действия:
 - бонусные игры, со сбором фишек/рыбок, которые можно в дальнейшем обменивать на скидки и/или льготное оформление второго продукта.
- клиентам пользующимся одним продуктом банка постараться продать еще один продукт, т.к. клиенты с двумя продуктами менее склонны к оттоку, предложения по второму продукту должны учитывать потребности клиентов

3 - уделить внимание контролю качества, удовлетворенности от использования продуктов

- разработка скрипта для обзона, собирающего информацию о удовлетворенности от использования и пожеланий клиентов, это позволит оптимизировать действующие и разработать новые продукты.
- организовать сбор обратной связи удовлетворенности клиентов из зоны риска.

Реализация всех вышеописанных рекомендаций требует тщательной проработки профильными подразделениями банка

CUSTOMER CHURN



Спасибо за внимание!