

Анализ оттока клиентов банка

Подготовил:
Сучков И. К.

Дата:
26.04.2023г.

Содержание

- Цель
- Общая информация
- Поиск взаимосвязи оттока клиентов.
- Портрет клиента для привлечения
- Сегментация
- Проверка гипотез
- Вывод

Цель исследования

- **Заказчик исследования:** отдел маркетинга Банка
- **Цель:** проанализировать клиентов банка и выделить сегменты клиентов, которые склонны оставаться в банке

Общая информация

Ссылка на дашборд:

- [Dashboard](#)

Источник данных:

- Файл - [bank_scrooge](#)

Данные для анализа:

- в нашем распоряжении имеется файл с информацией о клиентах банка в документе имеются
- 10 000 записей, содержащих следующую информацию:
- - `userid` — идентификатор пользователя,
- - `score` — баллы кредитного скоринга,
- - `City` — город,
- - `Gender` — пол,
- - `Age` — возраст,
- - `equity` — количество баллов собственности
- - `Balance` — баланс на счёте,
- - `Products` — количество продуктов, которыми пользуется клиент,
- - `CreditCard` — есть ли кредитная карта,
- - `last_activity` — активный клиент,
- - `estimated_salary` — заработная плата клиента,
- - `Churn` — ушёл или нет.

Какой отток клиентов в банке?

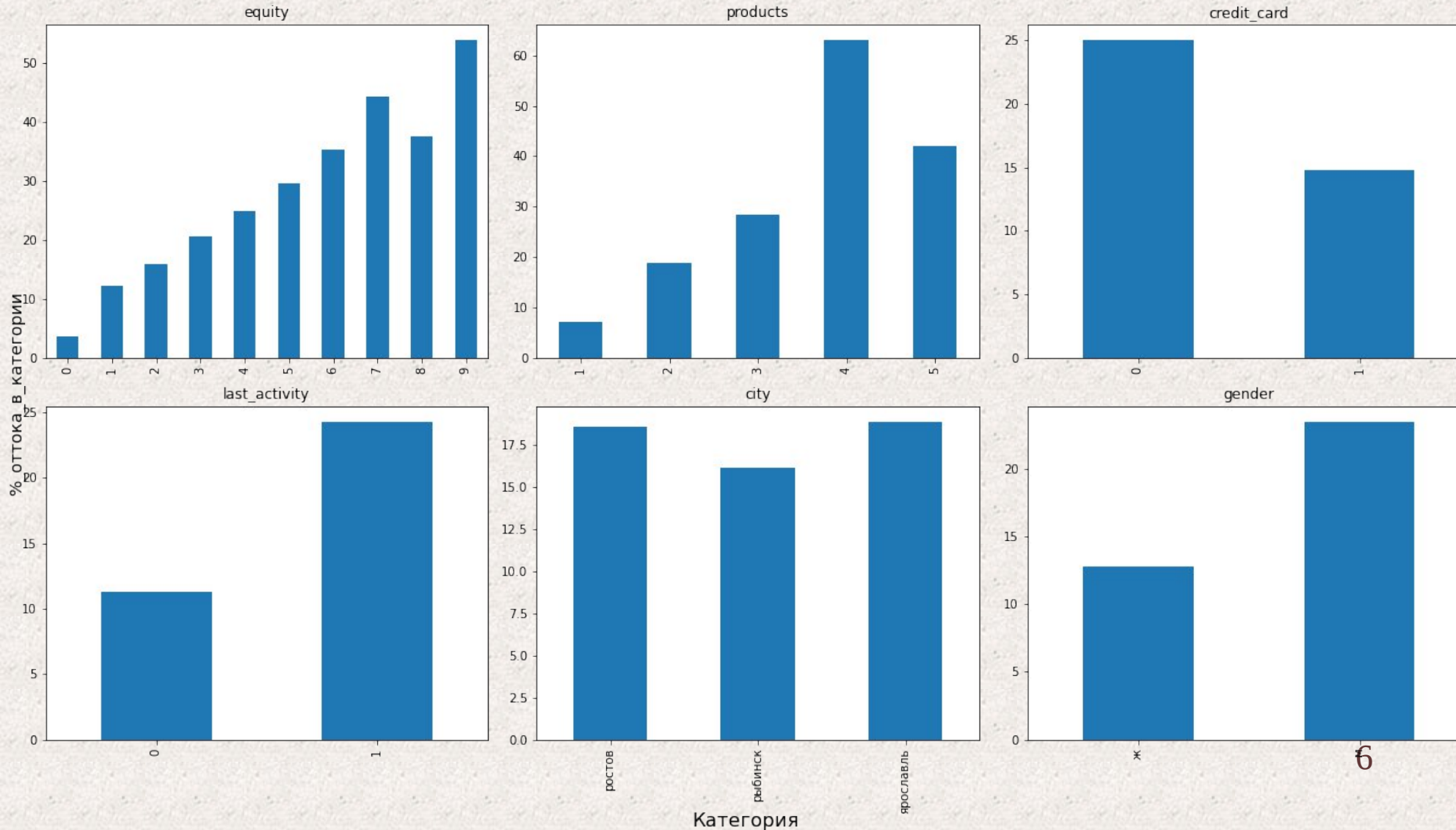
Соотношение клиентов, которые планируют уйти и остаться



- Отток клиентов в Банке составляет 18 процентов.

На основе чего составлялся портрет клиента?

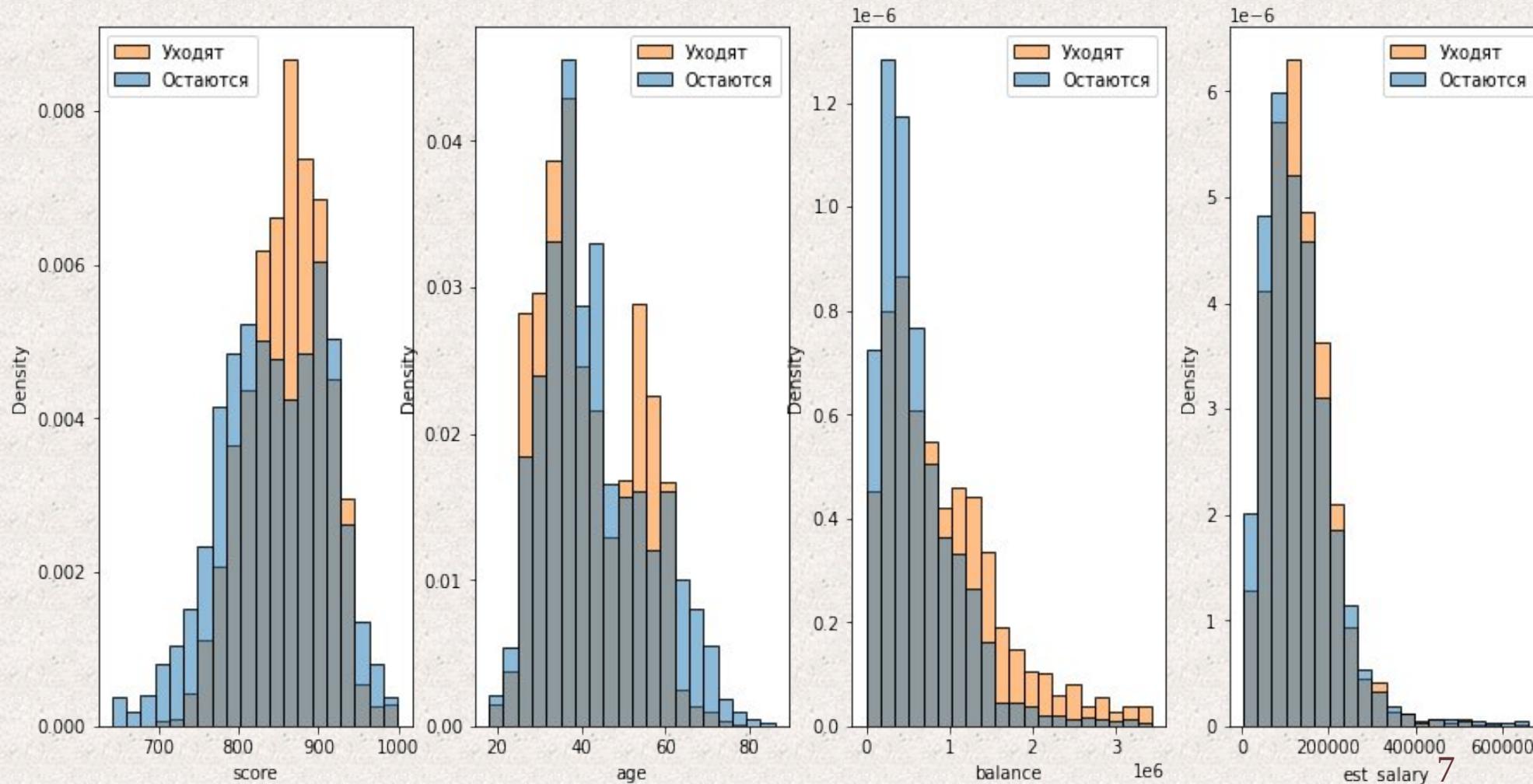
Отток клиентов из банка для категориальных переменных



- **Диапазон лояльных клиентов:**
- Оценка ≤ 2
- Продуктов ≤ 2 .
- Кредитная карта есть
- Активности не было
- Город Рыбинск
- Пол женский

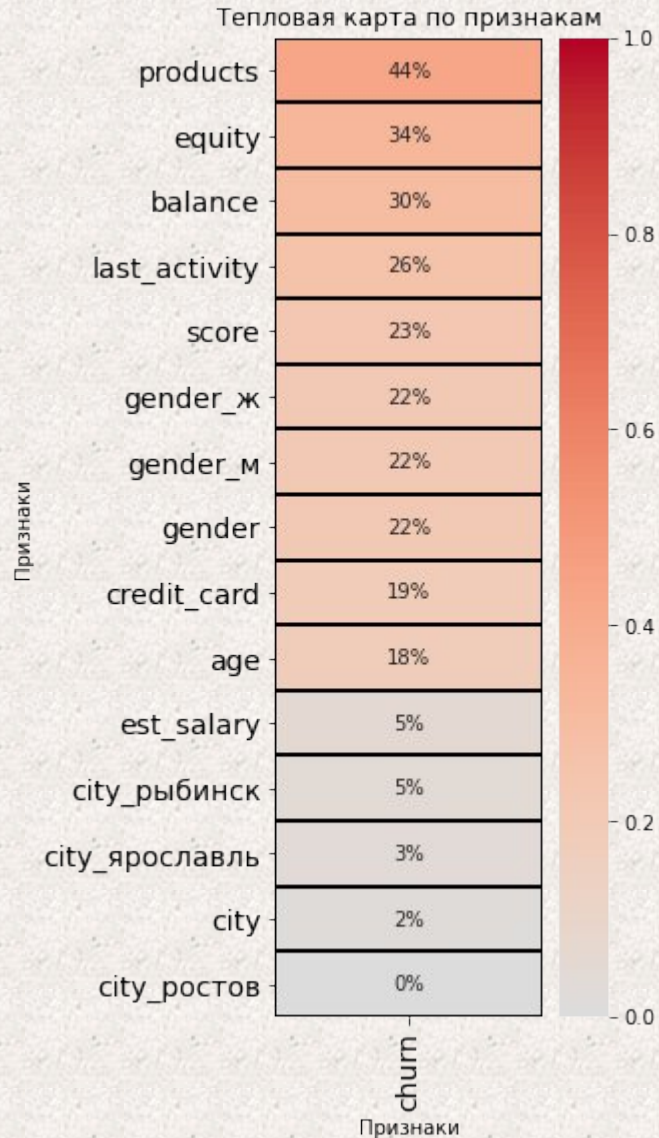
На основе чего составлялся портрет клиента?

Сопоставимые диаграммы



- **Диапазон лояльных клиентов:**
- Рейтинг до 820, 910-920, 940 и более
- Возраст до 24, 32, 39-45, 47 - 50, 61
- Баланс до 725, 850-900, 1050-1100 тыс.
- ЗП до 100, и от 230 тыс. и более

На основе чего составлялся портрет клиента?



- Наибольшая корреляция наблюдается с количеством продуктов 44%
- За тем оценка собственности 34%
- После баланс 30%

Портрет среднего лояльного клиента и клиента, который вероятней всего не уйдет.

- **Средний лояльный клиент:**

- Оценка собственности 2-3.
Ближе к 2.
- Баланс около 613 тыс.
- Активность необязательна
- Пол скорее женский
- Количество продуктов 1-2.
Ближе к 3.
- Город Ярославль
- Возраст 43
- ЗП около 130 тыс.
- Рейтинг 844

- **Диапазон лояльных клиентов:**

- **Важные показатели:**

- Оценка собственности 2 и менее
- Баланс до 725тыс., 850-900, 1050-1100.
- Женщина
- Количество продуктов от 2 и меньше.
- Кредитная карта есть

- **Менее важные:**

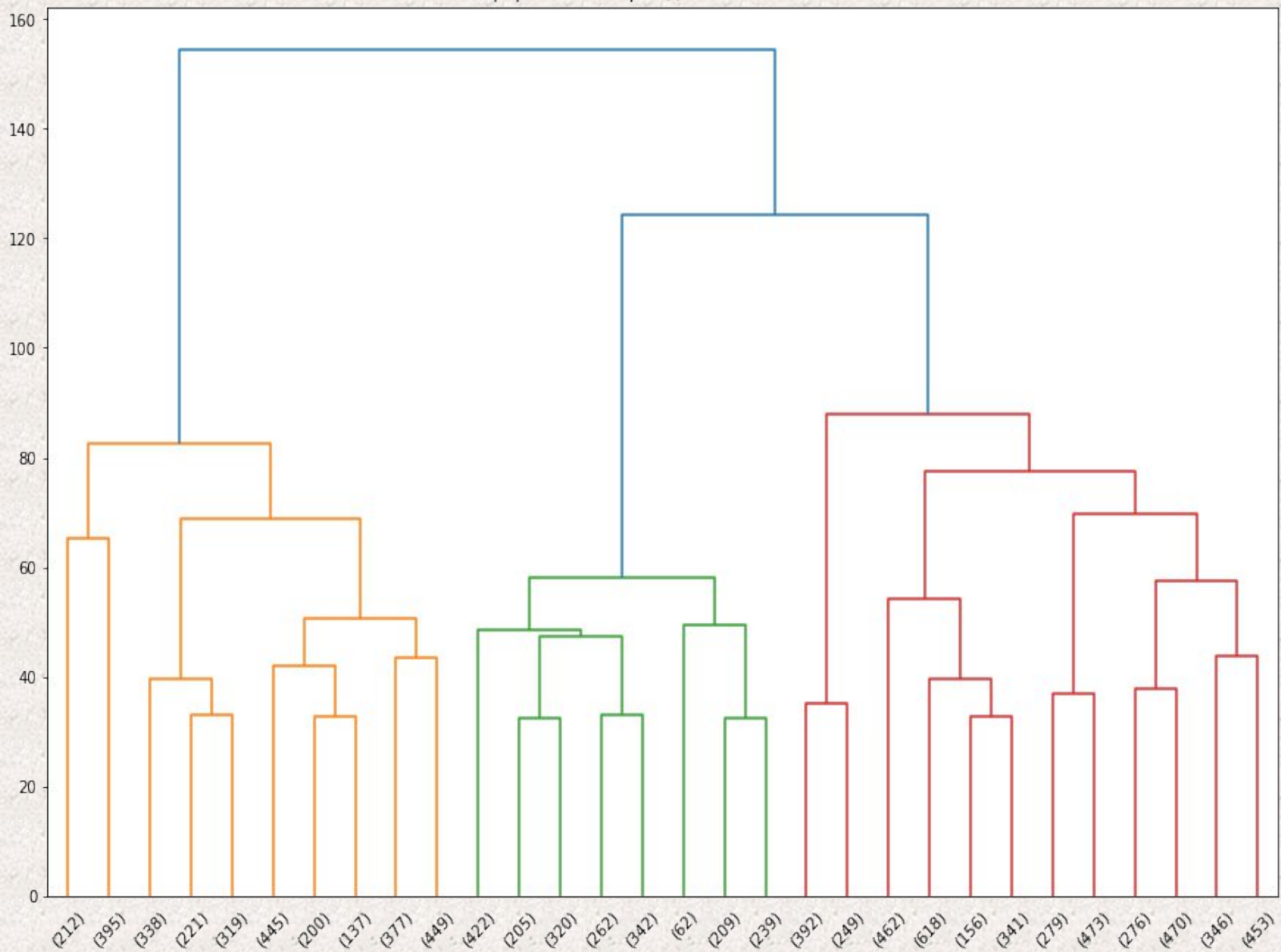
- города Рыбинск
- ЗП до 100000 и от 230000
- Возраст 24, 32, 39-45, 47 - 50, 61.
- рейтинг до 820, 910-920, от 940

Какая модель лучше может предсказать отток?

- **Построены модели прогнозирования оттока:**
 - логистической регрессией
 - случайным лесом
- **Метрики для модели логистической регрессии:**
 - Accuracy: 0.85
 - Precision: 0.64
 - Recall: 0.29
- **Метрики для модели случайного леса:**
 - Accuracy: 0.89
 - Precision: 0.79
 - Recall: 0.49
- **Лучший результат показала модель:**
 - Случайный лес

Кластеризация

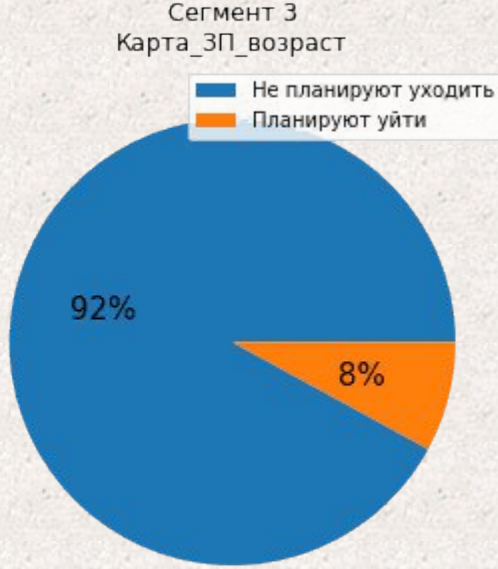
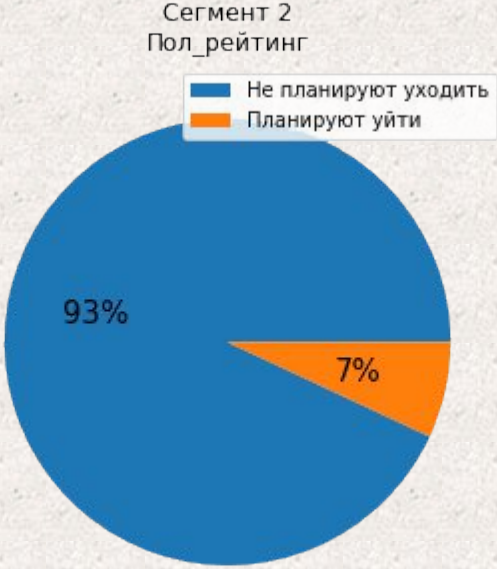
Иерархия кластеров для клиентов банка



- Для кластеризации использовался метод KMeans
- С помощью дендограммы выделено три кластера.
- Качество проведенной кластеризации не высокое Silhouette_score = 0,13,
- Средние значения характеристик по кластерам не противоречат и подтверждают сделанные ранее выводы о портретах клиентов.

cluster_km	0	1	2
score	852.700129	841.820330	849.208270
age	41.488380	42.188998	44.892281
equity	0.387011	3.771029	3.613809
balance	84825.384939	705102.848858	820389.793238
products	1.257586	2.052805	2.322838
credit_card	0.853131	1.000000	0.000379
last_activity	0.536798	0.493288	0.545144
est_salary	166471.272808	109894.218889	118423.938961
city	1.645255	1.851588	1.796282
gender	0.567463	0.521220	0.380501

Сегментация



- Оценка ценности

	Segment	credit_card_%	score	est_salary	equity	razmer
2	segment_3	100.0	836	83753	2.37	239
0	segment_1	75.7	853	130672	0.67	4255
1	segment_2	65.9	798	106142	2.35	2187

- Сегмент 1 - Высокий потенциал. 75% имеют кредитные карты, самый высокий рейтинг, ЗП и количество клиентов, но низкую оценку собственности.
- Сегмент 2 - Средний. Самое низкое держателей кредиток 66%, самый низкий рейтинг, средний уровень ЗП, но хорошая оценка недвижимости и средняя численность.
- Сегмент 3 - Наиболее стабильный. 100% используют кредитки, средний рейтинг, ниже всех ЗП, но хорошая оценка недвижимость, и при это самая мало численный сегмент.

Равен ли доход и возраст лояльных клиентов?

- Для ответа провели проверку статистических гипотез о равенстве средних двух выборок.
 - H_0 : Средний доход отточных и неотточных клиентов равен
 - H_1 : Средний доход отточных и неотточных клиентов различается
 - H_0 : Средний возраст отточных и неотточных клиентов равен
 - H_1 : Средний возраст отточных и неотточных клиентов различается
 - Статистический уровень значимости $\alpha = 5\%$
- Отвергаем нулевую гипотезу, считаем, что средний доход отточных и неотточных клиентов различаются.
- Есть основания отвергнуть нулевую гипотезу, считаем, что средний возраст отточных и неотточных клиентов различается.

Равен ли доход и возраст лояльных клиентов?

- Для ответа провели проверку статистических гипотез о равенстве средних двух выборок.
 - H_0 : Средний доход отточных и неотточных клиентов равен
 - H_1 : Средний доход отточных и неотточных клиентов различается
 - H_0 : Средний возраст отточных и неотточных клиентов равен
 - H_1 : Средний возраст отточных и неотточных клиентов различается
 - Статистический уровень значимости $\alpha = 5\%$
- Отвергаем нулевую гипотезу, считаем, что средний доход отточных и неотточных клиентов различаются.
- Есть основания отвергнуть нулевую гипотезу, считаем, что средний возраст отточных и неотточных клиентов различается.

Вывод

- **В данной работе были проанализированы данные о клиентах банка.**
- Построены модели прогнозирования оттока клиентов.
- Проверены две статистические гипотезы.
- Выделены основные черты клиентов, которых надо привлечь:
 - Важные показатели
 - Менее важные
- **Рекомендации:**
 - Если у человека небольшой показатель оценки недвижимости, то вероятней всего его заинтересует сниженная ставка на ипотеку. Рекомендация для сегмента 1.
 - Женщина важно спокойствие, поэтому им можно предложить какие-нибудь страховые услуги. Для всех сегментов.
 - Кредитные карты можно выдавать без платы за обслуживание. Но для этого нужно будет провести дополнительное исследование. Рекомендация для сегмента 2.
 - Рекламную компанию необходимо развернуть в городе Рыбинск. Для всех сегментов.
 - Для сегмента 3 можно предлагать такой продукт, как накопительный счет. Это поднимет баланс в сегменте.
- **Рекомендации отделу маркетинга:**
 - Обратить внимание в первую очередь на важные показатели. То есть необходимо будет проводить рекламную компанию среди клиентов женского пола, у которых оценка собственности 2 и менее.

Спасибо за внимание!