

Анализ оттока клиентов банка «Метанпром»

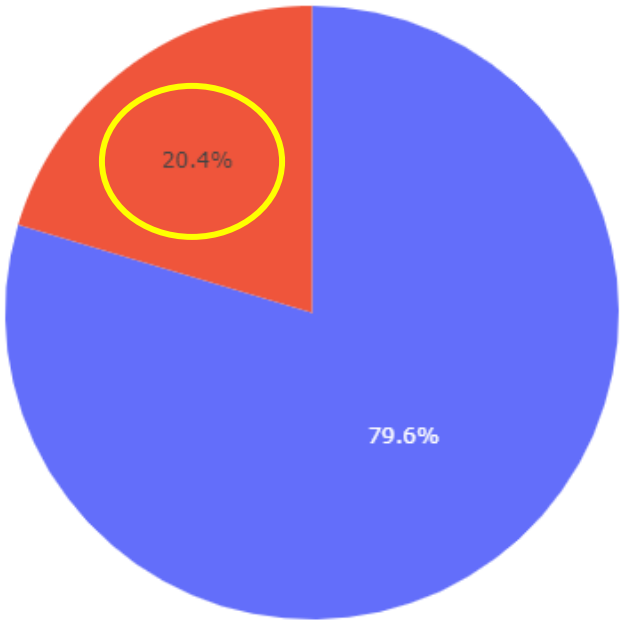
Описание датасета

Имеющиеся данные:

- `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,
- `estimated_salary` — заработная плата клиента,
- `Churn` — ушёл или нет.

Соотношение клиентов банка по оттоку

Соотношение клиентов банка по оттоку

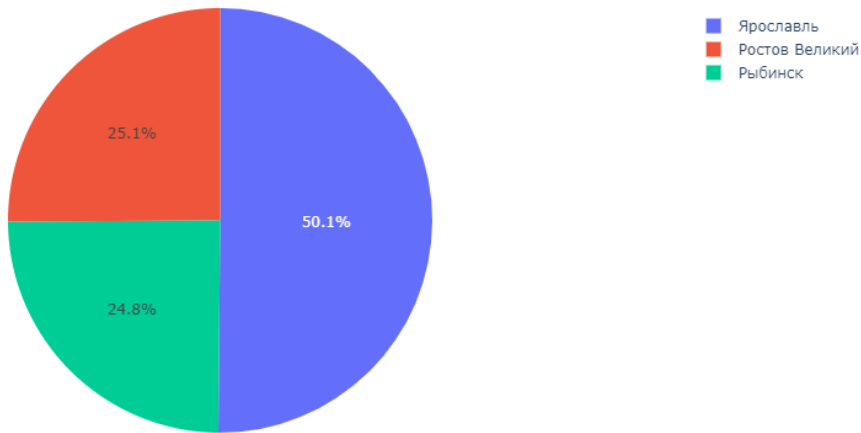


Ключевые моменты:

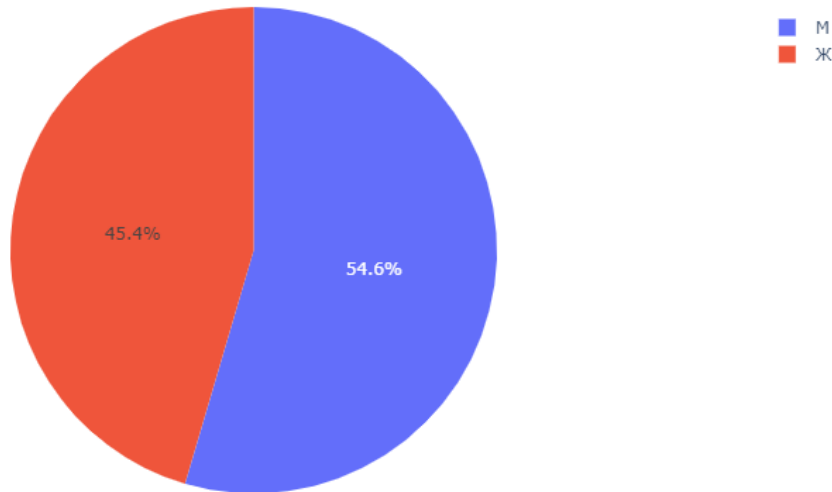
- В нашем распоряжении один датасет с данными о клиентах банка «Метанпром»
- 20% клиентов ушли в отток

Соотношение клиентов банка по полу и городу

Соотношение клиентов банка по городу



Соотношение клиентов банка по полу

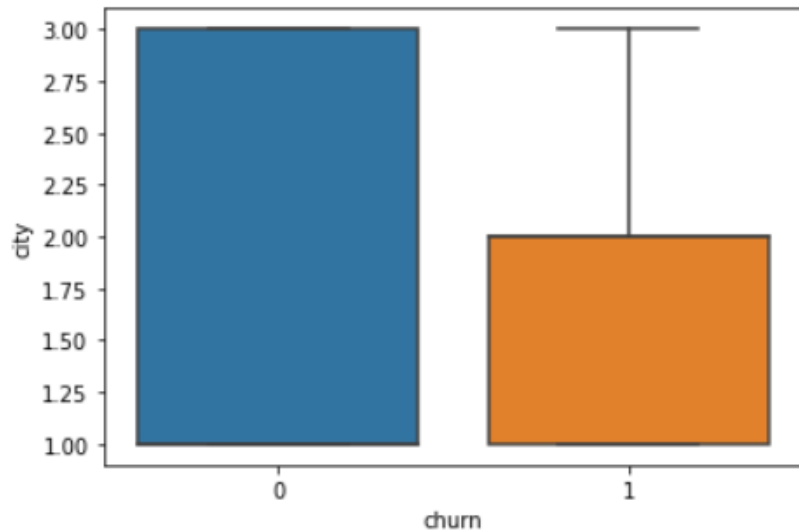


Ключевые моменты:

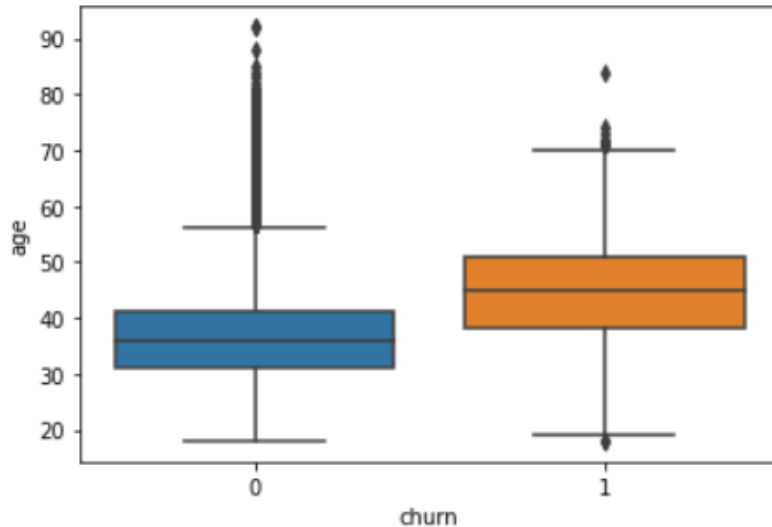
- Половина клиентов проживает в Ярославле, по 25% в Ростове Великом и Рыбинске
- Чуть больше половины (55%) – это мужчины

Портрет клиента в разрезе факта оттока

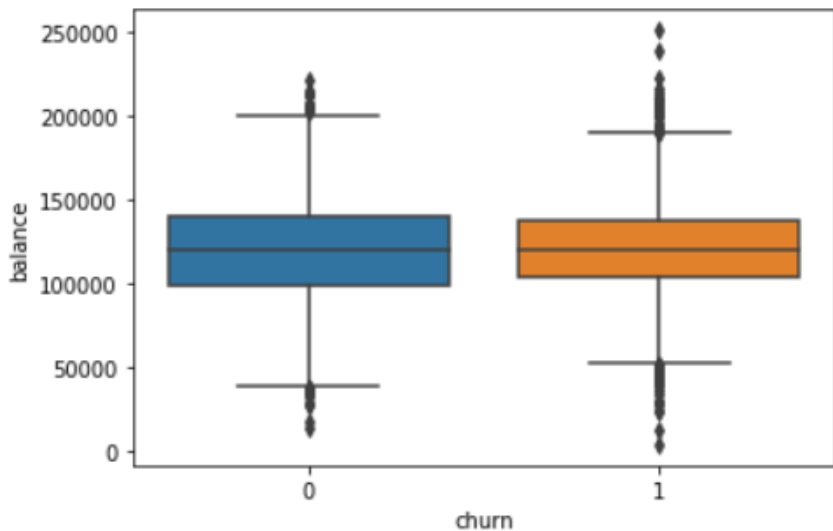
Город



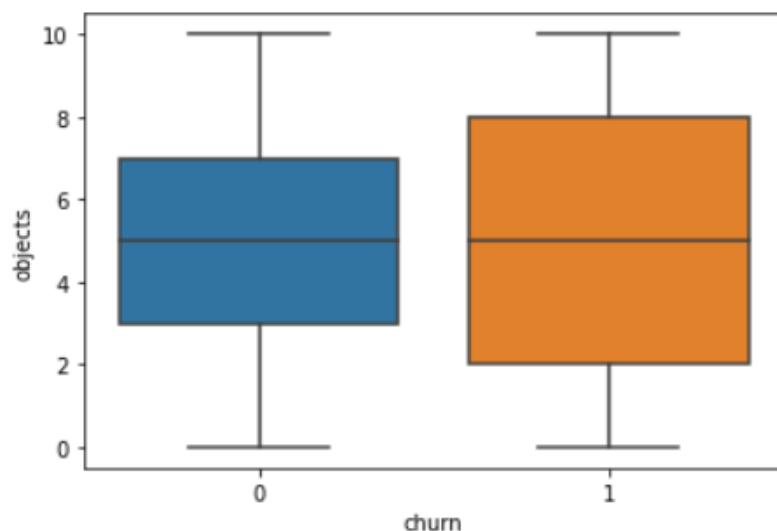
Возраст



Баланс



Количество объектов в собственности

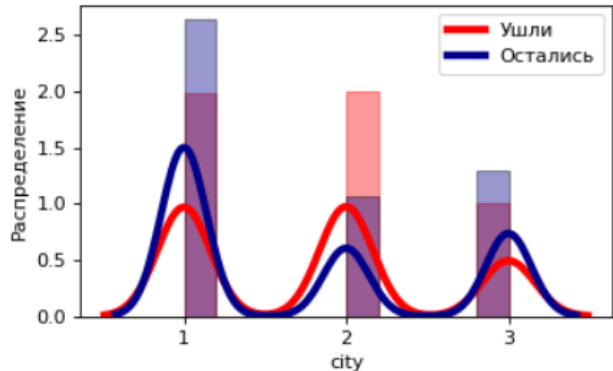


Ключевые моменты:

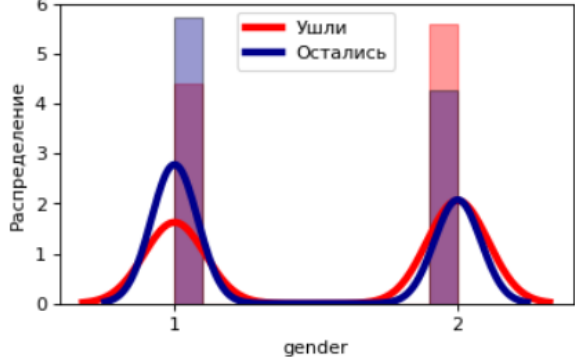
- чаще всего в отток уходят клиенты из Ярославля и Ростова Великого
- в возрасте 40-55
- владеющие недвижимостью в количестве от 2 до 8 объектов
- имеющие суммы на балансе чуть меньше, чем у оставшихся клиентов

Распределение клиентов

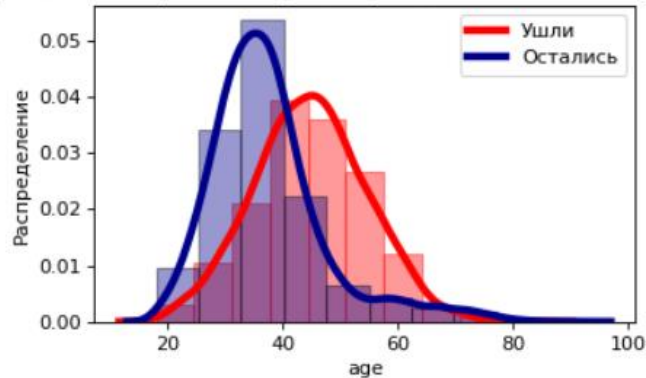
Распределение признака city для ушедших и оставшихся клиентов



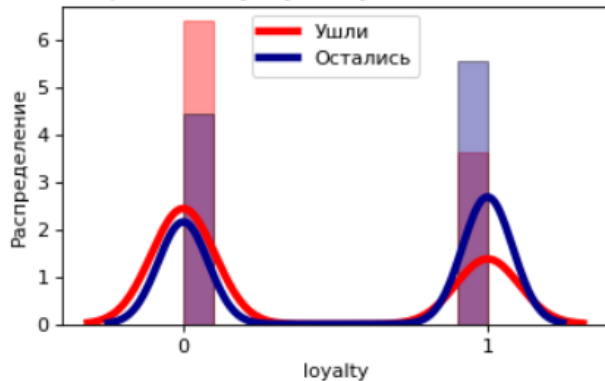
Распределение признака gender для ушедших и оставшихся клиентов



Распределение признака age для ушедших и оставшихся клиентов



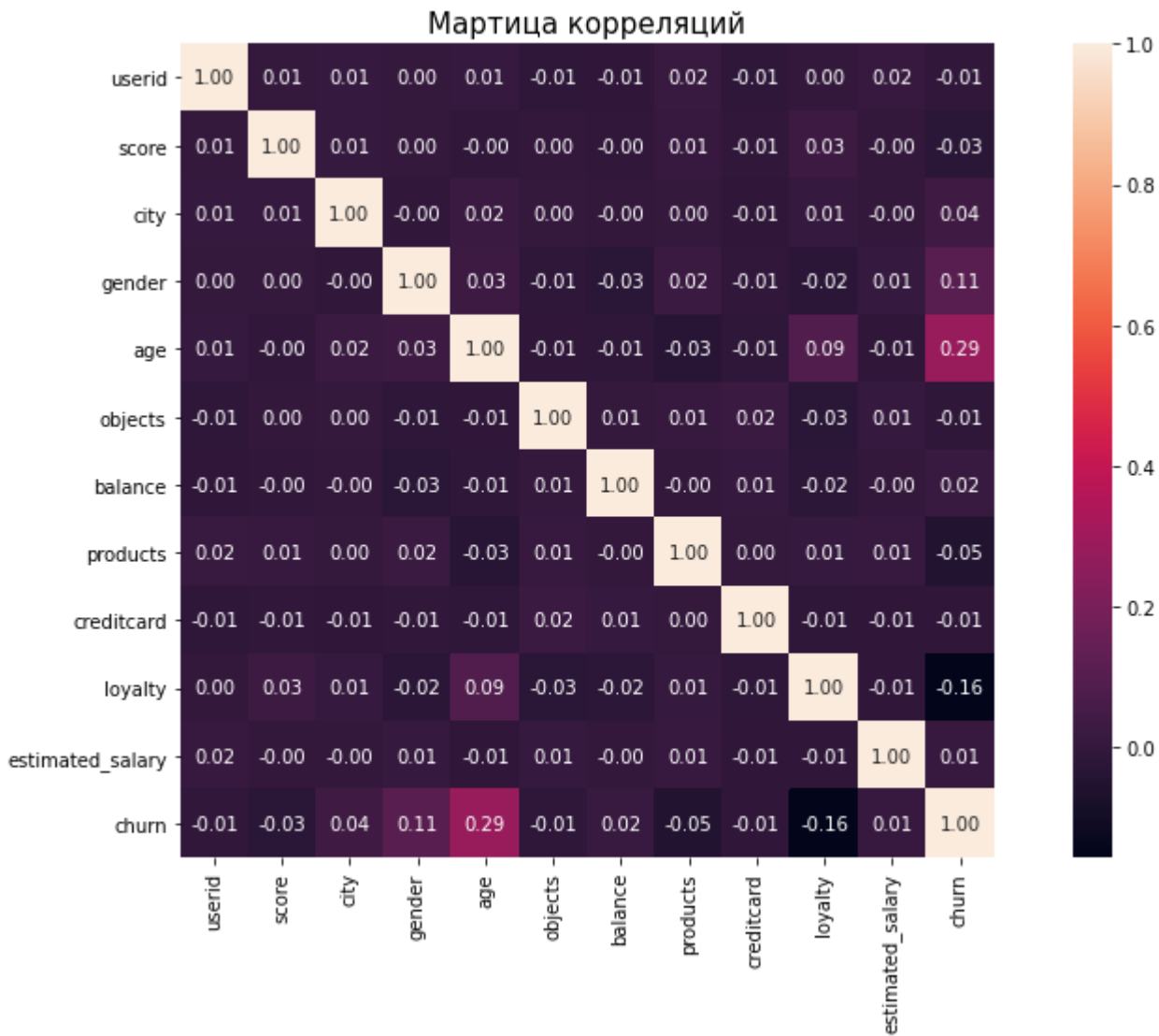
Распределение признака loyalty для ушедших и оставшихся клиентов



Ключевые моменты:

- Чаще всего уходят клиенты из Ростова Великого
- Женщины
- Возраст ушедших чуть больше возраста оставшихся
- Признак активности у ушедших отсутствует

Наличие корреляции



Ключевые моменты:
Больше всего между собой коррелируют возраст и факт оттока

Кластеризация на основании алгоритма K-Means

Ключевые моменты:

- По модели кластеризации больше всего склонны к оттоку 0, 1 и 3 кластеры. 2 кластер менее всего подвержены оттоку.

cluster	score	city	gender	age	objects	balance	products	creditcard	loyalty	estimated salary	churn
0	650.45	1.75	2.00	39.22	4.99	119,179.13	1.54	1.00	0.50	99,866.69	0.25
1	651.35	1.76	1.46	39.11	4.91	119,519.12	1.53	0.00	0.52	100,974.40	0.21
2	652.73	1.75	1.00	39.48	5.05	119,346.15	1.53	1.00	1.00	98,914.38	0.10
3	646.95	1.72	1.00	37.49	5.18	121,984.09	1.51	1.00	0.00	100,353.38	0.23

Описание кластеров

- Отличительные особенности кластера номер 0:
 - в него входят только женщины
 - имеют кредитную карту
 - каждая вторая - активный клиент
 - самая высокая доля оттока
- Противоположные ему кластер номер 2:
 - в него входят только мужчины
 - они так же имеют кредитную карту
 - абсолютно все из них активные клиенты
 - самая низкая доля оттока
- Кластер номер 1 выделяется отсутствием кредитной карты и смешанный составом (мужчины и женщины).
- Кластер номер 3:
 - чуть моложе остальных
 - состоит исключительно из мужчин
 - не имеет в себе активных клиентов
 - меньше всего продуктов на клиента
 - больше всего объектов в собственности

Рекомендации по удержанию клиентов

- Кластер 0:
 - т.к. это женщины, активные клиенты с кредитными картами, то рекомендуется добавить им программу лояльности. Например кэшбэк при оплате кредиткой в салонах красоты или попадание операций связанных с красотой и шопингом в грейс период.
- Кластер 1:
 - т.к. у представителей данного кластера отсутствуют кредитные карты, можно предложить им кредитку с бесплатным годовым обслуживанием.
- Кластер 3:
 - самый молодой кластер, имеющий больше всего объектов в собственности. Рекомендуется стимулировать активность клиентов предложив "мужскую" дебетовую карту с программой кэшбэка на категории, связанные со строительством, спортом, автомобилями.

Проверка гипотез

После проведения Ttest'a о равенстве среднего двух генеральных совокупностей по взятым из них выборкам при уровне значимости 0,05 были получены следующие результаты для выбранного уровня значимости:

- средняя заработная плата тех, кто ушел и тех, кто остался равны
- возраст тех, кто ушел и тех, кто остался не равны.