



PREDICTIVE  
TECHNOLOGIES



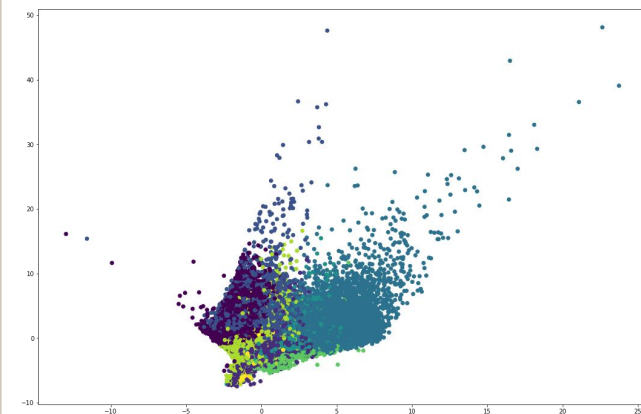
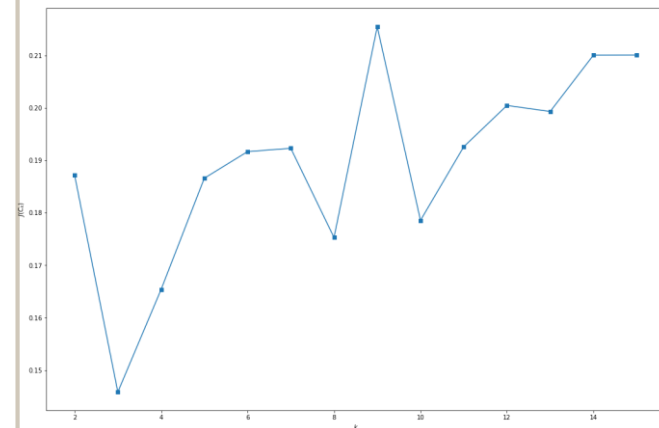
# Кластеризация клиентов ломбарда

24.09.18

# Методы кластеризации

## Основные предпосылки:

- 1** Использовалась суммарная ретроспектива активности клиента за все время, но рассматривались только клиенты, имевшие визит в ломбард с 01 июня 2016 по 01 июня 2018 года (275 783 клиента)
- 2** Кластеризация проводилась методом *k*-средних, который хорошо зарекомендовал себя как точный и быстрый алгоритм
- 3** Для отрисовки кластеров применялся метод главных компонент (PCA) для преобразования многомерного пространства в двумерное



# Методы кластеризации

## Используемые модели

### 1 Было проверено несколько гипотез:

- ▶ Кластеризация всех новых последних залоговых билетов клиентов
- ▶ Сегментация на 1 билет, от 2х до 57 и больше 57 залоговых билетов (25%, 70% и 5% соответственно)
- ▶ Группировка всех перезалогов в первый залоговый билет

### 2 Каждая из групп разбивалась на различное количество кластеров, для которых вычислялись коэффициенты силуэта и среднеквадратичные отклонения от центров – таким образом выбиралось оптимальное количество кластеров.

В результате была определена оптимальная структура данных для анализа – это выборка из всех последних новых залоговых билетов по клиентам, за исключением 1% - у которых их количество превышало 73.

# Кластеризация

## Данные

Кластеризация производилась по полям:

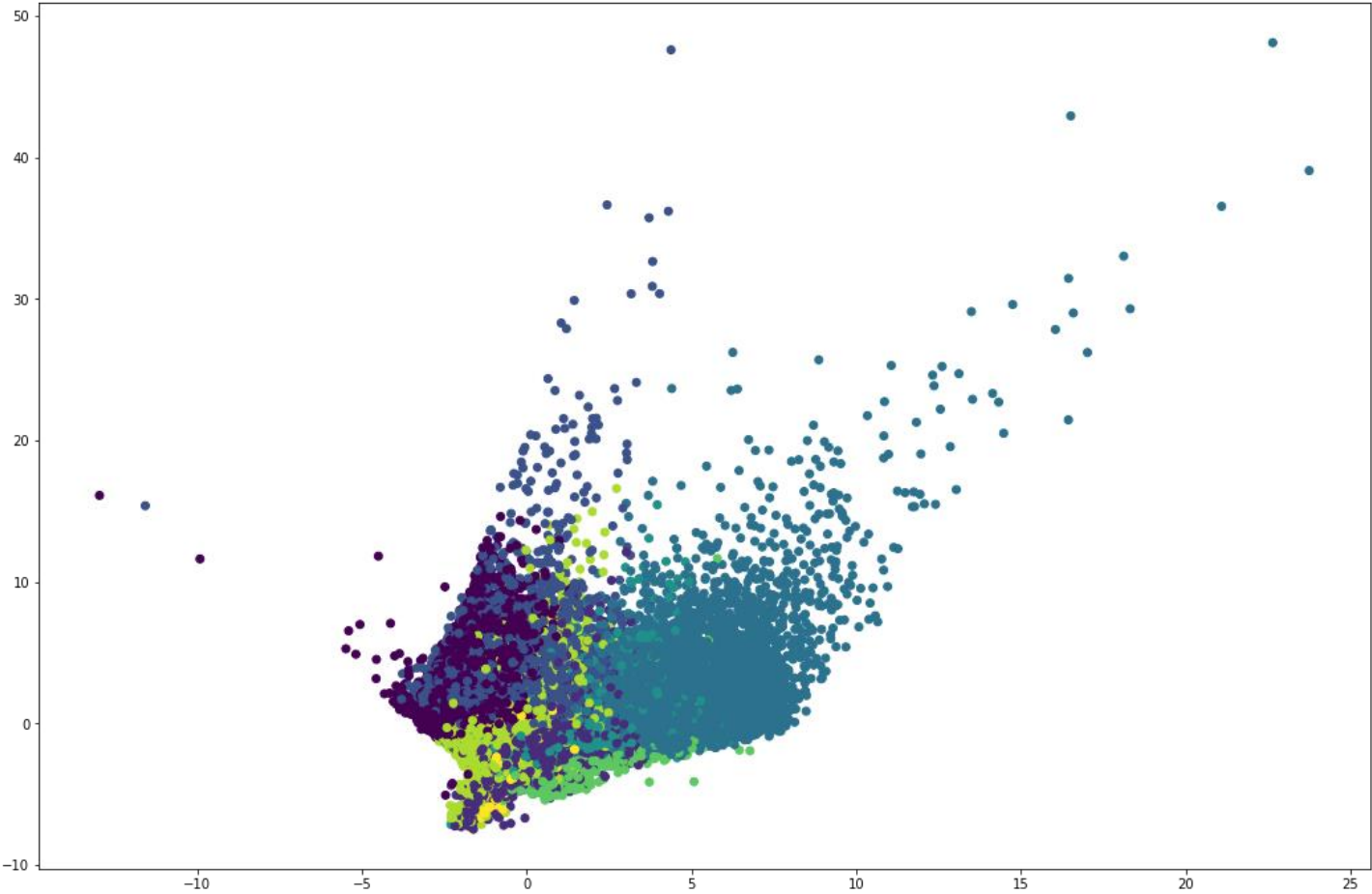
1. Возраст
2. Life Time
3. LTV
4. Средняя сумма займа
5. Количество визитов
6. Процент по СМС
7. Процент по промо
8. Процент новых
9. Процент перезалогов
10. Процент ОСН
11. Процент скупки
12. Дней залога
13. Дней с прошлого визита
14. Вес Gross
15. Цена за грамм
16. Рейтинг клиента

275783 - клиентов	Среднее	Среднее	Среднее	Среднее	Среднее	Среднее	Среднее	Среднее	Среднее
Кластер	1	2	3	4	5	6	7	8	9
Доля кластера	32.0%	9.1%	1.4%	16.3%	0.4%	16.9%	7.4%	15.4%	1.2%
Доля мужчин	44%	20%	61%	39%	28%	39%	37%	29%	40%
Возраст	34	53	41	34	34	35	36	41	35
Рейтинг клиента	0.0	0.0	1.4	1.0	6.0	2.2	3.0	3.3	7.0
Life Time	645	800	997	881	1240	1340	1491	1592	1851
Будет перезалог	16%	27%	37%	22%	8%	24%	37%	56%	25%
LTV	8513	28428	625588	29145	40997	44242	80575	210593	33801
Средняя сумма займа	7832	10793	74505	7326	6275	6192	6996	9826	6190
Количество визитов	1.1	3.2	10.3	4.0	6.5	7.3	11.7	22.2	5.4
Процент по СМС	0%	2%	2%	5%	2%	4%	2%	1%	1%
Процент по промо	1%	83%	17%	12%	19%	11%	7%	12%	13%
Процент возвратов	46	65	73	55	9	71	80	92	75
Процент новых	100%	93%	80%	89%	89%	88%	75%	47%	85%
Процент перезалогов	0%	5%	19%	5%	1%	10%	24%	53%	15%
Процент ОСН	0%	0%	0%	0%	2%	0%	0%	0%	0%
Процент скупки	0%	3%	1%	6%	8%	2%	1%	0%	0%
Дней залога	15	24	31	24	10	28	42	53	33
Дней с прошлого визита	3	72	49	155	127	219	116	66	517
Вес Gross	6.36	8.85	63	6.16	5.60	5.72	6.64	9.73	7.06
Цена за грамм	1291	1268	1263	1248	1124	1169	1146	1111	960
Средняя себестоимость	8254	11313	79030	7764	6332	6768	7661	10789	6673
Максимальная себестоимость	8430	13782	120268	11745	13370	11811	14740	23148	11677
Средний процент в день	35%	24%	28%	35%	36%	35%	34%	32%	31%
Минимальный процент в день	35%	23%	25%	33%	31%	30%	29%	25%	24%
Максимальный процент в день	35%	27%	32%	37%	40%	40%	39%	38%	36%
Этот ЗБ будет возвращен	46%	64%	65%	51%	21%	62%	66%	78%	62%
Будет визит в течении 3 месяцев	19%	32%	43%	30%	15%	30%	44%	63%	26%

## Визуализация

### Вклад компонент в оси XY

	PC-1	PC-2
Age	0.075054	0.168748
LifeTime	0.281675	-0.233881
SumAmountzb	0.209939	0.269559
AVGAmountzb	0.048371	0.468608
ZB	0.325913	0.004105
PercentSMS	0.016833	-0.061949
PercentPromo	-0.026223	0.138188
PercentBudetPerezalog	0.269219	0.151148
PercentGood	0.253700	0.059101
PercentNew	-0.338174	-0.027763
PercentPerez	0.372064	0.055321
PercentOSN	-0.004543	-0.054893
PercentBuyUp	-0.032021	-0.030911
AVGDneyZaloga	0.299169	0.022758
AVGDneySproshZB	0.033221	-0.293959
AVGWeightGross	0.079117	0.456145
AVGCostPerGram	-0.147330	0.189275
AVGPercentPerDay	-0.050576	-0.211829
MoreGood	0.330763	-0.092912
SEX_0	0.052446	-0.091384
SEX_1	-0.052446	0.091384
CreditCondition_0	-0.247163	0.296743
CreditCondition_1	-0.025351	-0.115663
CreditCondition_2	0.067614	-0.218517
CreditCondition_3	0.098101	-0.071845
CreditCondition_4	0.159610	0.021157
CreditCondition_5	0.178465	0.028175
CreditCondition_6	-0.015548	-0.029999
CreditCondition_7	0.025011	-0.097268
CreditCondition_8	0.024194	0.022383



# Кластеризация

## Описание

### 1 Кластер 1 – 32%.

Кластер клиентов, который воспользовались услугами ломбарда всего 1 раз. Это люди в основном с 1 визитом и всегда это Залог. СМС им не отправляли, в промо не участвуют. Процент возвратов залога меньше половины, 16% вернется на перезалог, еще 3% придут повторно в течении 3х месяцев. Самая высокая цена за грамм, при стандартном проценте в день - 0.35%.

### 2 Кластер 2 – 9.1%

Преимущественно женщины (80%) пенсионного и предпенсионного возраста. Имеют по 3 и больше визита, самую большую сумму среднего займа и самую минимальную ставку 0,24%. 83% приходят по промо – самый большой процент. В основном это новый залог (93%) из которых 22% придут на перезалог. На СМС реакция слабая. Периодичность визитов – раз в 2-3 месяца.

### 3 Кластер 3 – 1.3%

Небольшой, но очень значимый кластер «профессионалов». Средняя сумма залога в 10 раз больше обычной. В 61% это мужчины около 40 лет. В среднем первый визит был 3 года назад, количество визитов 10. Есть реакция на промо. Берут в основном на месяц и в 75% возвращают.

### 4 Кластер 4 – 16.3%

Кластер тех, кто из первых 2х перерос и стал ходить в ломбард регулярно. Здесь уже по 4 билета в среднем, но процент возврата еще на уровне 55%. Как следствие рейтинг не вырос сильно - 1 и процент в день довольно высок. Люди редко перезаклаываются. Частота визитов примерно раз в пол года. Хорошая реакция на СМС - 5% и небольшая на промо 12%. Высок процент скупки

### 5 Кластер 5 – 0.4%

Скупщики - небольшой процент людей с высоким рейтингом (6), самым высоким процентом в день (36%) и низкими возвратами. Здесь самый высокий процент скупки (8%). Каждый 5й визит по промо (19%), реагирует на СМС только 2%. Периодичность визитов - 3 раза в год. Процент возврата залога – 9% - минимальный. Сомнительный кластер – непонятно откуда высокий рейтинг при таком возврате.

# Кластеризация

## Описание

### 6 Кластер 6 – 16.9%.

Клиенты, имеющие уже по 7-8 визитов. Из них 10% это перезалоги. В среднем погашают задолженность за месяц. Частота визитов - раз в 7-8 месяцев, при этом через 3 месяца вернется четверть, многие для перезалога. Есть неплохой отклик на СМС - 4% и небольшой на промо 11%.

### 7 Кластер 7 – 7.4%

Постоянные клиенты. В среднем первый раз приходили 4 года назад. За это время у них было более 11 визитов. Процент новых займов с увеличением количества визитов начинает падать, а процент перезалогов расти. Клиенты, которые понимают что будут пользоваться услугами ломбарда регулярно предпочитают возвращать займ, чтобы уменьшать процентную ставку. Реакция на промо и СМС минимальна.

### 8 Кластер 8 – 15.4%

Почти 15% базы - клиенты которые ходят в ломбард регулярно - раз в 2 месяца. Первый визит был 4.5 года назад и за это время они приходили более 22 раз. Самый высокий процент возвратов (92%), при том что перезалогов (53%) уже больше чем оформления новых (47%). Приносят уже либо большие, либо несколько вещей – самый большой вес. Реакция на СМС минимальна – 1%, и небольшая на промо – 12%, при этом скупкой или ОСН не интересуются.

### 9 Кластер 9 – 1.2%

1% клиентов, это самые старые клиенты, посещающие ломбард крайне редко - раз в полтора - два года. У них невысокая средняя сумма займа, неплохой процент возвратов. И достаточно низкие ставки, за счет высокого рейтинга. На СМС практически не реагируют, реакция на промо - 13%. Вызывает вопросы – возможно клиенты искусственно повысили себе рейтинг.

# Схематичное представление кластеров

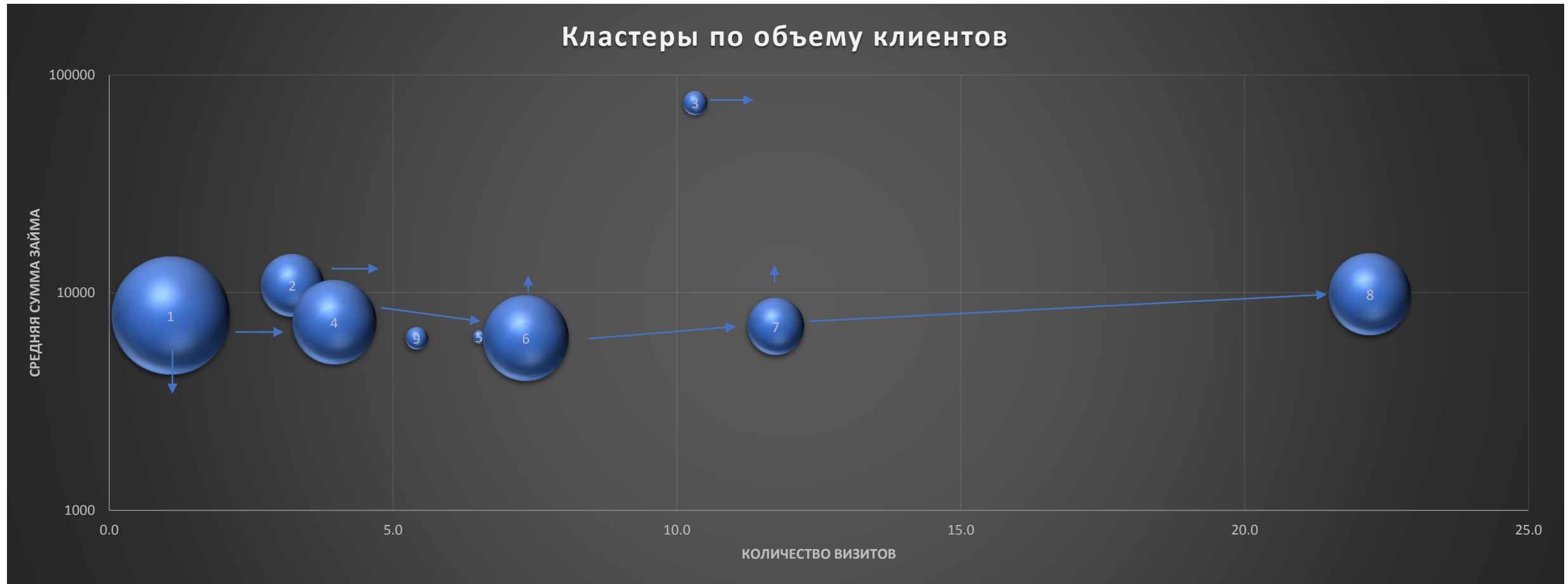
LTV кластера от количества покупок и средней суммы чека





# Схематичное представление кластеров

## Распределение клиентов



# Рекомендации

- **Кластер 1** – ходят за деньгами, пользуются высокой стоимостью оценки за грамм и в половине случаев не возвращают. Необходимо стимулировать прийти второй раз – возможно при возврате первого займа, второй оценивается так же высоко. Промо на повторный визит.
- **Кластер 2** – самые возрастные и пользующиеся промо, их нужно выводить на большее количество визитов. Продолжать промо акции, на которые они реагируют. Процент поднимать не стоит, необходимо стимулировать ходить чаще.
- **Кластер 3** – профессионалы, сами знают когда приходить. Пользуются промо, видимо для снижения процентной ставки, рейтинг не высок. Что бы не потерять, можно чуть снизить процентную ставку.
- **Кластер 4** – молодые люди – эволюция первого кластера. Есть реакция на смс. Этот кластер нужно стремиться переводить в 6й кластер, но с большей суммой. Работать над лояльностью.
- **Кластер 5** – по факту это скупщики, можно привлекать при необходимости золота для переработки.
- **Кластер 6** – большой кластер, необходима стимуляция к увеличению суммы ЗБ. Можно увеличить стоимость за грамм или стимулировать ходить чаще. Плавно перетекают в 7й кластер. Работать над лояльностью.
- **Кластер 7** – видно увеличение средней суммы по сравнению с 6м кластером. Возможно сказывается уменьшение процента. Если сравнивать с 8 кластером, то прослеживается прямая зависимость средней суммы залога от процента в день. Реакция на смс слабая, но остается реакция на различные промо.
- **Кластер 8** – постоянные клиенты, задача их не отпустить. Реакции на СМС нет. Так же можно уменьшить процент или штраф за просрочку, для стимуляции продолжать обращаться.
- **Кластер 9** – необходимо проверить, откуда высокий рейтинг при небольшом количестве визитов. Необходимо либо снижать рейтинг, либо повышать процент (UPD. 6-й рейтинг является негативным, плохие клиенты).
- **Эволюция** клиентов: 1=>4=>6=>7=>8. Ядро клиентов – кластеры 6, 7 и 8. Кластеры 2 и 3 – обособленные, провоцировать к увеличению количества или частоты визитов.