

Анализ клиентов фитнес-центров «Культурист-датасаентист»



Цель работы: сеть фитнес-центров «Культурист-датасаентист» разрабатывает стратегию взаимодействия с клиентами на основе аналитических данных. Необходимо провести анализ и подготовить план действий по удержанию клиентов.

Выполнил Олег Дмитриев

Учебные данные предоставил Яндекс.Практикум

Наши клиенты

У **50%** клиентов
длительность абонемента
1 месяц

50% клиентов
посещают зал не чаще
2-х раз в неделю

50% клиентов являются
сотрудниками компаний-
партнёров клуба

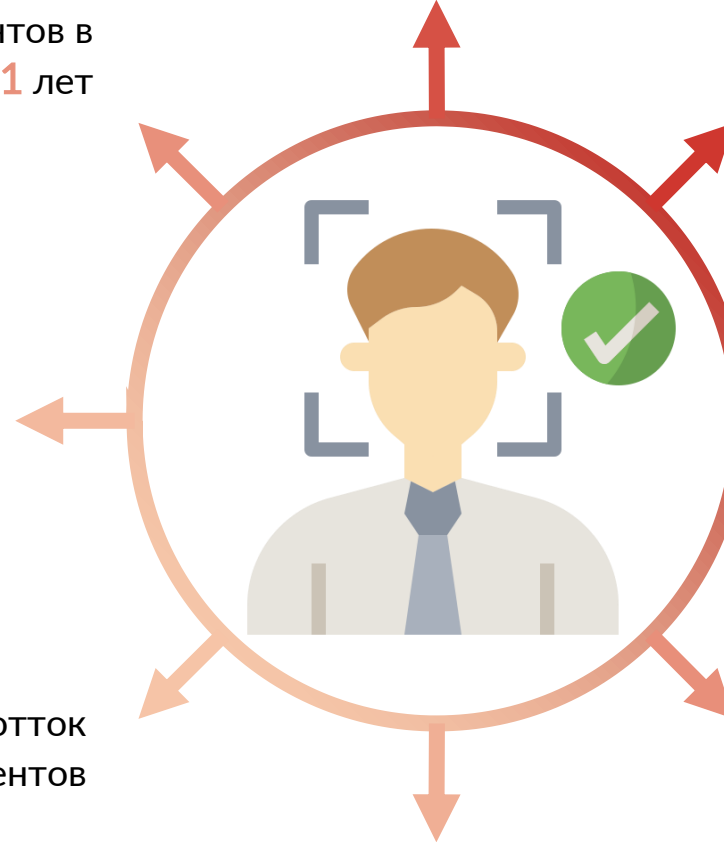
30% клиентов
пришли в рамках
акции «приведи
друга»

40% клиентов ходят на
групповые программы
тренировок

В текущем месяце в отток
попало **26%** клиентов

Половина клиентов не превысила
лайфтайм в **3 месяца**;
3/4 клиентов не превысило лайфтайм в
5 месяцев с начала занятий

Большинство клиентов в
возрасте **27-31** лет



Выявленные закономерности для ушедших в отток клиентов



- доля ушедших в отток клиентов больше среди тех, кто не является партнером компании;
- большая часть клиентов, попавших в отток, имела контракт на 1 месяц;
- доля ушедших в отток клиентов больше среди тех, кто не посещал групповые занятия;



- больше всего в отток уходит клиентов в возрасте 22-30. Клиенты, которые остались в среднем старше (27-35);
- срок до окончания текущего действующего абонеента для клиентов, ушедших в отток, в подавляющем большинстве случаев составляет 1 месяц;



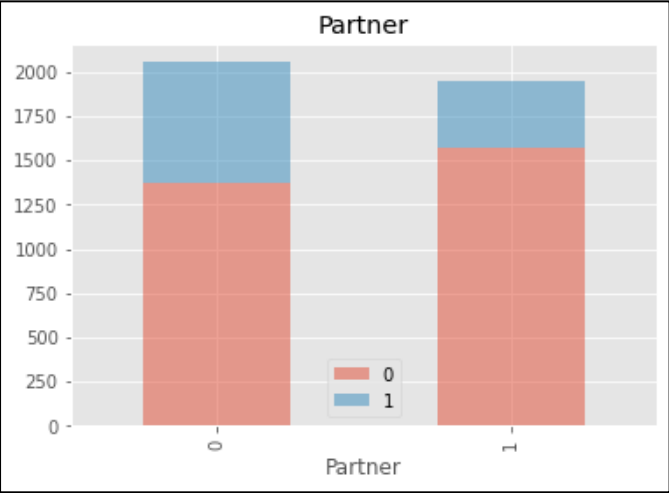
- самая большая вероятность ухода в отток среди тех клиентов, кто занимается менее 3-х месяцев (среднее время формирования привычки для человека);



- средняя частота посещений в неделю за все время с начала действия абонеента для тех, кто ушел в отток в основном составляет 1-2 раза. Для тех, кто посещает зал 3 и более раз - вероятность оттока значительно меньше.

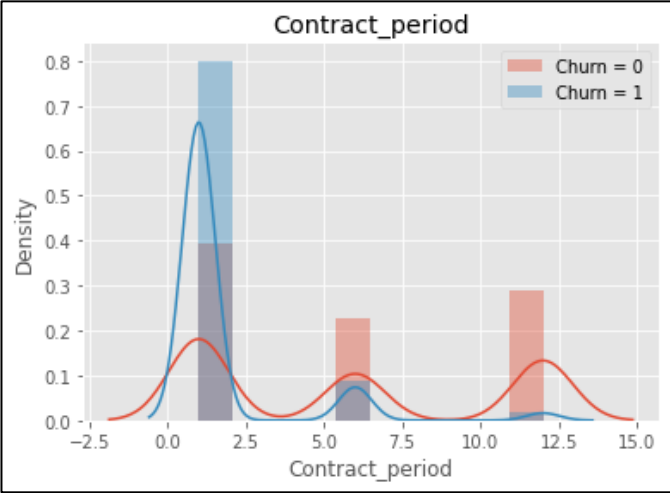
Сравнение распределений признаков для ушедших и оставшихся клиентов

Кол-во клиентов ушедших в отток среди тех, кто ушел в отток и тех, кто остался



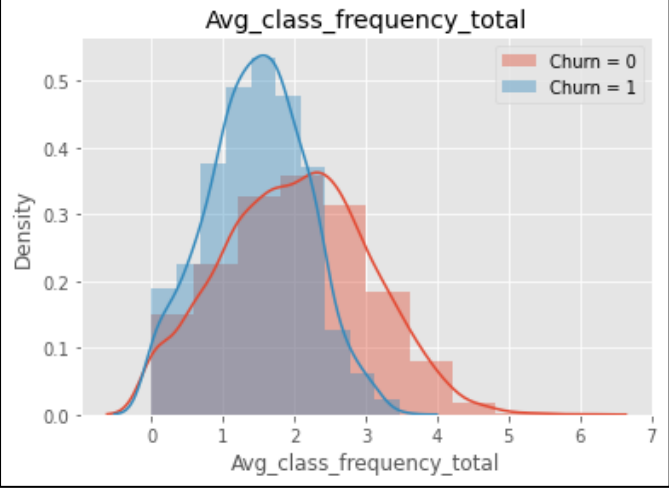
- ушли в отток
- остались

Кол-во клиентов ушедших в отток среди тех, кто ушел в отток и тех, кто остался



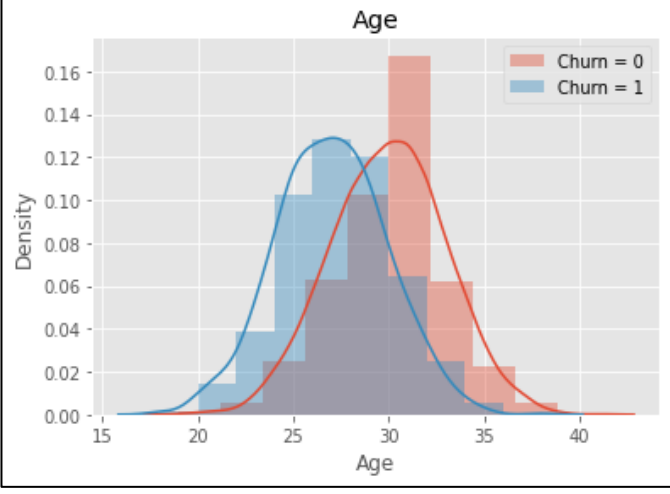
- ушли в отток
- остались

Кол-во клиентов ушедших в отток среди тех, кто ушел в отток и тех, кто остался



- ушли в отток
- остались

Кол-во клиентов ушедших в отток среди тех, кто ушел в отток и тех, кто остался



- ушли в отток
- остались

Матрица корреляции признаков:

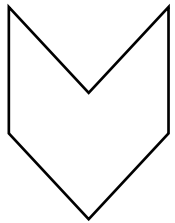
Значимые взаимосвязи между признаком оттока **Churn** и:

- **Lifetime** — время с момента первого обращения в фитнес-центр (в месяцах);
- **Avg_class_frequency_current_month** — средняя частота посещений в неделю за предыдущий месяц;
- **Age** — возраст;
- **Contract_period** — длительность текущего действующего абонеента (месяц, 6 месяцев, год);
- **Month_to_end_contract** — срок до окончания текущего действующего абонеента (в месяцах).

gender	1	0.0067	-0.01	-0.005	-0.0085	0.0005	0.018	0.014	-0.0093	-0.0013	0.014	0.015	0.0092	0.00071
Near_Location	0.0067	1	-0.0051	0.21	-0.016	0.15	0.15	0.058	0.041	0.14	0.071	0.043	0.063	-0.13
Partner	-0.01	-0.0051	1	0.45	0.01	0.31	0.023	0.047	0.023	0.29	0.061	0.025	0.046	-0.16
Promo_friends	-0.005	0.21	0.45	1	0.002	0.24	0.12	0.05	0.037	0.24	0.073	0.028	0.054	-0.16
Phone	-0.0085	-0.016	0.01	0.002	1	-0.0069	-0.01	-0.011	0.0093	-0.011	-0.019	0.0083	0.013	-0.0012
Contract_period	0.0005	0.15	0.31	0.24	-0.0069	1	0.17	0.14	0.11	0.97	0.17	0.096	0.16	-0.39
Group_visits	0.018	0.15	0.023	0.12	-0.01	0.17	1	0.081	0.032	0.16	0.076	0.057	0.076	-0.18
Age	0.014	0.058	0.047	0.05	-0.011	0.14	0.081	1	0.08	0.13	0.16	0.12	0.18	-0.4
Avg_additional_charges_total	-0.0093	0.041	0.023	0.037	0.0093	0.11	0.032	0.08	1	0.11	0.079	0.044	0.08	-0.2
Month_to_end_contract	-0.0013	0.14	0.29	0.24	-0.011	0.97	0.16	0.13	0.11	1	0.16	0.085	0.15	-0.38
Lifetime	0.014	0.071	0.061	0.073	-0.019	0.17	0.076	0.16	0.079	0.16	1	0.11	0.18	-0.44
Avg_class_frequency_total	0.015	0.043	0.025	0.028	0.0083	0.096	0.057	0.12	0.044	0.085	0.11	1	0.95	-0.25
Avg_class_frequency_current_month	0.0092	0.063	0.046	0.054	0.013	0.16	0.076	0.18	0.08	0.15	0.18	0.95	1	-0.41
Churn	0.00071	-0.13	-0.16	-0.16	-0.0012	-0.39	-0.18	-0.4	-0.2	-0.38	-0.44	-0.25	-0.41	1
gender		Near_Location	Partner	Promo_friends	Phone	Contract_period	Group_visits	Age	Avg_additional_charges_total	Month_to_end_contract	Lifetime	Avg_class_frequency_total	Avg_class_frequency_current_month	Churn

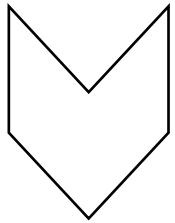
Построение модели прогнозирования оттока клиентов

Построение модели бинарной классификации клиентов, где целевой признак — факт оттока клиента в следующем месяце



Обучение моделей:

Модели были обучены по двум алгоритмам "Logistic Regression" и "Random Forest"



Сравнение моделей по метрикам accuracy, precision и recall:

Модель Logistic Regression показала себя лучше в сравнении с Random Forest по метрике Recall (полнота).

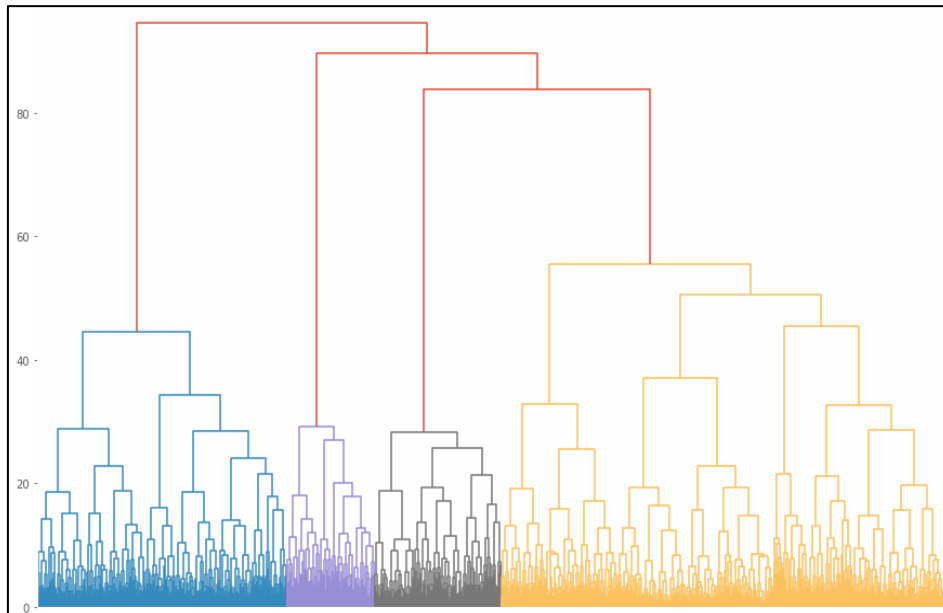
Метрика Precision (точность) выше у алгоритма Random Forest

Logistic Regression	
Accuracy	0.90
Precision	0.79
Recall	0.81

Random Forest	
Accuracy	0.90
Precision	0.80
Recall	0.77

Кластеризация клиентов

Способ №1: построение матрицы расстояний



С помощью методов `dendrogram`, `linkage` из `scipy.cluster.hierarchy` установлено оптимальное число кластеров 4 — четыре разных цвета на нижнем уровне дендрограммы

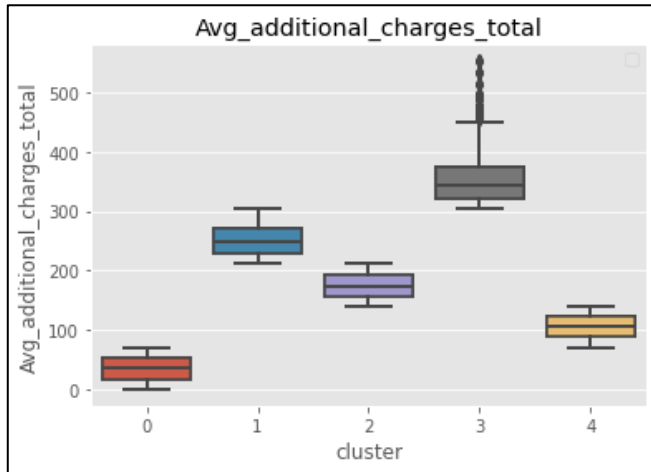
Способ №2: алгоритм K-Means

Номер кластера	Отток в кластере
0	0.350536
1	0.178470
2	0.227743
3	0.044280
4	0.333010

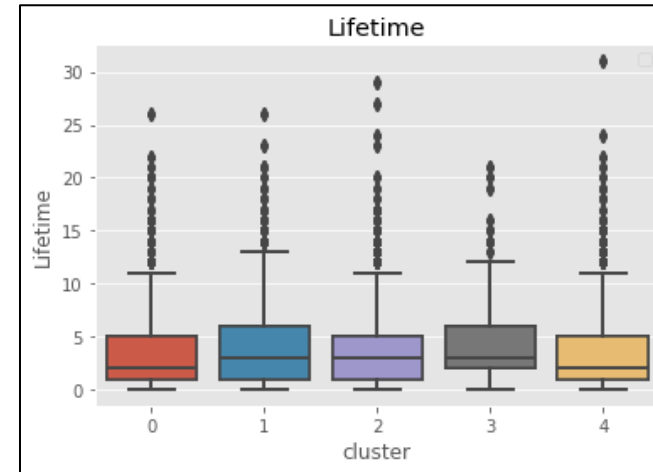
Алгоритм `K-Means` из библиотеки `sklearn.cluster` разделил данные по 5 кластерам. Больше всего отток в 0 и 4 кластерах:

Распределения признаков для разных кластеров

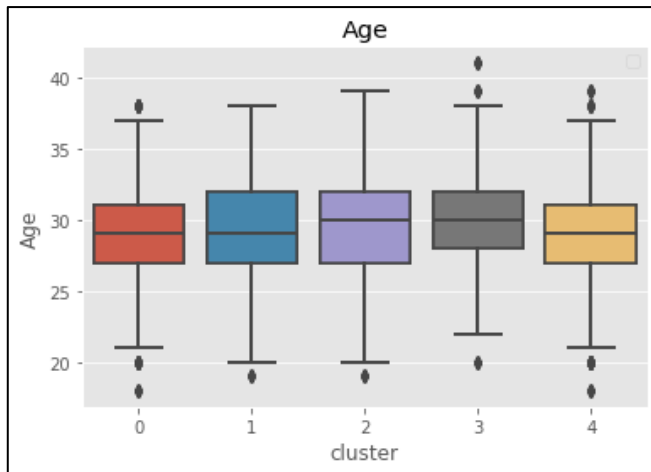
Распределение суммарной выручки с клиентов от дополнительных услуг фитнес-центра по кластерам



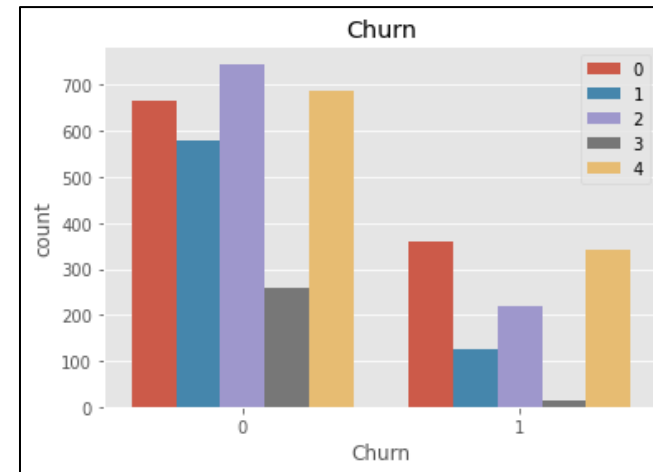
Распределение лайфтама клиентов по кластерам



Распределение возрастов клиентов по кластерам



Распределение оттока клиентов по кластерам (Churn 1 – ушли в отток)



Описание кластера №0

Клиенты нулевого кластера чаще всего попадают в отток. Люди из этого кластера поровну мужчины и женщины, чаще всего находятся рядом с залом (квартира/работа в одном районе). Половина из этих клиентов пришли по партнерской программе и 1/3 по акции "приведи друга". Для 90% клиентов известен телефон и 38% посещает групповые тренировки. Клиенты из 0 и 4 кластера самые молодые (большинство моложе 30) и они меньше всего пользуются дополнительными услугами фитнес-центра. Выручка в 0 кластере около 30 единиц с каждого клиента. Для большинства из них абонемент заканчивается через месяц. Большая часть клиентов не перешла лайфтайм в 3 месяца и посещает тренировки не чаще 2-х раз в неделю.

Описание кластера №1

Кластер имеет средний показатель оттока. Клиенты из этого кластера поровну мужчины и женщины, чаще всего находятся рядом с залом (квартира/работа в одном районе). Половина из этих клиентов пришли по партнерской программе и 1/3 по акции "приведи друга". Для 90% клиентов известен телефон и 43% посещает групповые тренировки. Клиенты из 1 кластера в большинстве моложе 30 и они пользуются дополнительными услугами фитнес-центра: кафе, спорттовары, косметический и массажный салон. Средняя выручка с этих клиентов около 250 единиц. Для большинства из них абонемент заканчивается через 5 месяцев. Большая часть клиентов перешла лайфтайм в 3 месяца и посещает тренировки не чаще 2-х раз в неделю.

Описание кластера №2

Кластер имеет средний показатель оттока. Клиенты из этого кластера поровну мужчины и женщины чаще всего находятся рядом с залом (квартира/работа в одном районе). Половина из этих клиентов пришли по партнерской программе и 1/3 по акции "приведи друга". Для 90% клиентов известен телефон и 42% посещает групповые тренировки. Клиенты из 2 кластера моложе 30 и они чаще всех пользуются дополнительными услугами фитнес-центра: кафе, спорттовары, косметический и массажный салон. Средняя выручка с этих клиентов около 330 единиц. Для большинства из них абонемент заканчивается через 1 месяц. Большая часть клиентов перешла лайфтайм в 3 месяца и посещает тренировки чаще 2-х раз в неделю.

Описание кластера №3

Кластер имеет самый низкий показатель оттока. Клиенты из этого кластера поровну мужчины и женщины, чаще всего находятся рядом с залом (квартира/работа в одном районе). Половина из этих клиентов пришли по партнерской программе и 1/3 по акции "приведи друга". Для 90% клиентов известен телефон и 43% посещает групповые тренировки. Клиентам из 3 кластера около 30 и они пользуются дополнительными услугами фитнес-центра: кафе, спорттовары, косметический и массажный салон. Средняя выручка с этих клиентов около 250 единиц. Для большинства из них абонемент заканчивается через 5 месяцев; Большая часть клиентов перешла лайфтайм в 3 месяца и посещает тренировки не чаще 2-х раз в неделю.

Описание кластера №4

Кластер имеет высокий показатель оттока. Люди из этого кластера поровну мужчины и женщины, чаще всего находятся рядом с залом (квартира/работа в одном районе). Половина из этих клиентов пришли по партнерской программе и 1/3 по акции "приведи друга". Для 90% клиентов известен телефон и 41% посещает групповые тренировки. Клиенты из 0 и 4 кластера самые молодые (большинство моложе 30) и они меньше всего пользуются дополнительными услугами фитнес-центра: кафе, спорттовары, косметический и массажный салон. С клиентов из 4 кластера выручка составляет около 100 единиц. Для большинства из них абонемент заканчивается через месяц. Большая часть клиентов не перешла лайфтайм в 3 месяца и посещает тренировки не чаще 2-х раз в неделю.

Выводы

1

Явные отличия среди кластеров выражены на графике Avg_additional_charges_total. Клиенты из 0 и 4 кластеров меньше всего пользуются дополнительными услугами фитнес-центра: кафе, спорттовары, косметический и массажный салон;

2

Значения признака Month_to_end_contract для 0 и 4 кластеров чаще всего составляют 1 месяц;

3

Для 0 и 4 кластеров Lifetime достигает пика на интервале 0-3 месяца.



Рекомендации по уменьшению оттока клиентов



Делать ставку на привлечение клиентов к дополнительными услугами фитнес-центра: кафе, спорттовары, косметический и массажный салон;



Привлекать клиентов 27-35 лет, так как они более надежные и платежеспособные;



Проявлять наибольшее внимание к клиентам, чей лайфтайм еще не превысил 3 месяца - вероятность их оттока самая большая;



Увеличивать число недельных посещений тренировок. Если клиент посещает тренировки более 2-х раз в неделю, то вероятность его оттока значительно меньше.

Спасибо за внимание!

Исследование выполнено в учебных
целях на основании данных,
предоставленных компанией
Яндекс.Практикум



Вопросы вы можете задать мне
по почте: doleg1999@yandex.ru
в telegram: [oleg_dmitriev_123](https://www.instagram.com/oleg_dmitriev_123)
Олег Дмитриев