

УДК 004.9

5.2.2. Математические, статистические и инструментальные методы в экономике

КЛАСТЕРИЗАЦИЯ КЛИЕНТСКОЙ БАЗЫ ОРГАНИЗАЦИИ С ИСПОЛЬЗОВАНИЕМ СПЕЦИАЛИЗИРОВАННОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ НА ОСНОВЕ МЕТОДА К-СРЕДНИХ

Фомина Елена Евгеньевна
к.т.н., доцент кафедры «Информатики и прикладной математики»
SPIN-код: 6602-8570, AuthorID: 642460,
f-elena2008@yandex.ru
Федеральное государственное бюджетное образовательное учреждение высшего образования «Тверской государственный технический университет», Тверь, Россия

Эффективное взаимодействие с клиентами – важная задача организации, которая требует использования проактивного и персонализированного подхода. Изучение поведения потребителей невозможно без использования современных инструментов анализа данных, к которым относится метод кластерного анализа. Его применение позволяет разделить аудиторию на группы на основе анализа взаимосвязей между характеристиками клиентов, которые имеют разнородный характер. Применение кластерного анализа невозможно без использования специализированного программного обеспечения. В статье описан программный комплекс, реализующий полную процедуру кластеризации клиентской базы методом к-средних, включающий в себя модули предобработки данных, определения оптимального числа кластеров, модули кластеризации и оценки качества проведенного анализа. Продемонстрированы возможности программного комплекса при анализе клиентской базы маркетплейса. Программный комплекс показал эффективность кластерного подхода для последующей оптимизации логистических и маркетинговых стратегий

Ключевые слова: КЛАСТЕРИЗАЦИЯ КЛИЕНТСКОЙ БАЗЫ, МЕТОД К-СРЕДНИХ, МЕТОД ЛОКТЯ, СИЛУЭТНЫЙ КОЭФФИЦИЕНТ, ПРИЛОЖЕНИЕ ДЛЯ КЛАСТЕРИЗАЦИИ КЛИЕНТСКОЙ БАЗЫ

<http://dx.doi.org/10.21515/1990-4665-216-032>

UDC 004.9

5.2.2. Mathematical, statistical and instrumental methods in economics

CLUSTERING OF THE ORGANIZATION'S CUSTOMER BASE USING SPECIALIZED SOFTWARE BASED ON THE K-MEANS METHOD

Fomina Elena Evgenievna
Candidate of technical sciences, associate professor,
Department of Computer Science and Applied Mathematics
RSCI SPIN-code: 6602-8570, AuthorID: 642460,
f-elena2008@yandex.ru
Federal State Budgetary Educational Institution of Higher Education «Tver State Technical University», Tver, Russia

Effective customer interaction is an important task for an organization, which requires a proactive and personalized approach. Studying consumer behavior is impossible without using modern data analysis tools, such as cluster analysis. This method allows to divide audience into groups based on the analysis of relationships between customer characteristics that are diverse. The application of cluster analysis is impossible without the use of specialized software. The article describes a software package that implements a complete procedure for clustering a customer base using the k-means method, including data preprocessing modules, modules for determining the optimal number of clusters, and modules for clustering and evaluating the quality of the analysis. The capabilities of the software package are demonstrated in the analysis of a marketplace customer base. The software package has shown the effectiveness of the cluster approach for subsequent optimization of logistics and marketing strategies

Keywords: CLUSTERING OF CUSTOMER DATABASE, K-MEANS METHOD, ELBOW METHOD, SILHOUETTE COEFFICIENT, SOFTWARE PACKAGE

Введение

Клиентская база – это не просто список покупателей, а один из основных стратегических ресурсов любой коммерческой организации, от которого зависит её развитие и процветание. Динамичность рынка, приводящая к возрастающей конкуренции, диктует необходимость бизнесу выстраивать с клиентами долгосрочные отношения [1]. Для увеличения прибыли, упрочнения позиций на рынке и обеспечения устойчивого развития компании требуется грамотное управление взаимоотношениями с клиентами. Это управление должно быть направлено на решение двух основных задач: привлечение новых клиентов, а также расширение и упрочнение взаимодействий с теми, кто уже входит в активную клиентскую базу.

Эффективное взаимодействие с текущими клиентами – важная стратегическая задача, которая требует ряда мероприятий:

формирование персонализированной коммуникации, т.е. клиенты должны получать актуальную информацию о новых товарах, услугах и комплексных предложениях с учетом их истории покупок, предпочтений и жизненного цикла [2];

использование гибкой и привлекательной системы скидок, бонусов и программ лояльности;

непрерывное улучшение сервисов взаимодействия с клиентами.

Анализ клиентской базы и формирование эффективных стратегий взаимодействия с различными группами целевой аудитории требуют применения специализированных инструментов, в том числе современных методов Data Mining. В частности, важным этапом такого анализа выступает деление клиентской базы на группы, которые позволят точнее определить потребности каждой категории контрагентов и оптимизировать маркетинговую активность. Для решения этой задачи используется два подхода: сегментация и кластеризация.

Сегментация – методология, основанная на анализе объективно

измеряемых характеристик, например, таких, как частота покупок, сумма потраченных клиентом средств, демографические признаки, географические критерии.

К основным методам сегментации клиентской базы относятся ABC/XYZ-анализ и RFM-анализ.

Суть ABC-анализа заключается в разделении всей совокупности клиентов на три категории в зависимости от финансового вклада, который они вносят в прибыль компании. Категория А – это наиболее ценные клиенты, которые вносят большой вклад в прибыль и не требуют больших ресурсов организации (20% клиентов, приносящих 80% прибыли); категория В – группа клиентов, составляющих «золотую середину» и обладающие потенциалом роста (30% клиентов, приносящие 15% прибыли); категория С – малоактивные клиенты, которые вносят незначительный вклад в прибыль компании (80% клиентов, которые приносят 5% прибыли) [3]. Выявление активных клиентов и клиентов с потенциалом роста позволит адресно направить усилия менеджеров для получения большей отдачи.

Суть XYZ-сегментации клиентской базы заключается в разделении клиентов на группы в зависимости от покупательской активности за выбранный временной период, продолжительность которого определяется целями исследования. Категория Х - клиенты с коэффициентом вариативности менее 10%, характеризуются стабильной величиной потребления и высокой степенью прогнозирования (покупают часто и постоянно, например, каждый месяц); Y - клиенты с коэффициентом вариативности от 10% до 25%, их потребление обусловлено сезонными колебаниями, что снижает возможности прогнозирования продаж; Z - клиенты с коэффициентом вариативности выше 25%, покупают нерегулярно и непредсказуемо, прогнозировать их действия практически невозможно [3].

Метод ABC/XYZ-анализ комбинирует два варианта анализа в результате чего клиентская база разделяется на девять групп (табл. 1).

Таблица 1. Разделение клиентской базы методом ABC/XYZ-анализа

| Сегмент | X | Y | Z |
|---------|--|--|---|
| A | Большая доля прибыли, стабильный спрос | Большая доля прибыли, колеблющийся спрос | Большая доля прибыли, непредсказуемый спрос |
| B | Средние объёмы прибыли, стабильный спрос | Средние объёмы прибыли, колеблющийся спрос | Средние объёмы прибыли, непредсказуемый спрос |
| C | Малозначительная прибыль, стабильный спрос | Малозначительная прибыль, колеблющийся спрос | Малозначительная прибыль, непредсказуемый спрос |

Метод RFM-анализа сегментации клиентов предполагает разделение клиентской базы на группы на основе трёх ключевых показателей: recency (давность сделки), frequency (частота сделки), monetary (потраченная сумма) и присвоение каждому клиенту ранга по каждому показателю (например, 1-3-2, т.е. активность-лояльность-вклад в выручку). Далее клиенты объединяются в группы в зависимости от итогового набора рангов и разрабатываются стратегии взаимодействия с каждой группой.

Преимущества сегментации заключаются в её простоте реализации и возможности оперативного внедрения полученных результатов в бизнес-процессы. Методы сегментации реализованы в большинстве современных CRM и ERP-системах для принятия обоснованных решений [4]. Однако недостатком этого подхода является отсутствие гибкости, учет только ограниченного ряда факторов, а не их комплексного взаимодействия.

Кластеризация – методология, основанная на выявлении скрытых закономерностей и сходств характеристик клиентов путем анализа больших массивов первичной информации, в которых каждый клиент характеризуется набором разнородных показателей. В отличие от процедуры сегментации при кластеризации критерии формирования групп определяются автоматически алгоритмами машинного обучения, что

обеспечивает гибкость и адаптивность модели, а также выявление скрытых факторов и паттернов поведения клиентов. Кластеризация позволяет выявлять неявные взаимосвязи между параметрами клиентов, обеспечивая глубокую детализацию их потребностей. Это даёт компаниям возможность создавать индивидуальные стратегии взаимодействия с клиентами, повышая их лояльность и увеличивая доходность бизнеса.

Кластеризация клиентской базы имеет не только важное практическое значение для организации, но и позволяет разрабатывать новые бизнес-модели и стратегии, которые являются основой для цифровой трансформации бизнеса. Как отмечает Е.В. Ярошенко [5] «комплексное персонифицированное обслуживание клиентов, оптимизация взаимодействия с ними и улучшение доступа клиентов в обособленной цифровой среде приводит к появлению революционных цифровых бизнес-моделей».

К алгоритмам кластеризации, которые используются для анализа клиентской базы относятся нейронные сети Кохонена, кластеризация методом k-средних, иерархическая кластеризация, метод DBSCAN, метод AID, дискриминантный анализ, метод PRIZM, методы классификации с помощью деревьев решений, методы гибкого и компонентного сегментирования [6-9].

По сравнению с алгоритмами сегментации, реализация алгоритмов кластеризации требует наличия специального программного обеспечения, основанного на технологиях машинного обучения. Кроме того, интерпретация результатов кластеризации может быть сложнее по сравнению с классической сегментацией, так как выделенные группы могут содержать большее количество уникальных признаков и требовать дополнительного анализа для разработки конкретных мер воздействия. Тем не менее ценность и потенциал данного подхода анализа клиентской базы достаточно высоки, так как он позволяет реализовывать целенаправленные коммуникации с клиентами.

Актуальной является задача разработки специализированного программного обеспечения, позволяющего реализовывать полный цикл кластеризации клиентской аудитории, включая этапы определения оптимального числа кластеров и оценки качества результатов кластеризации.

Описание программного комплекса

Для реализации полного цикла процесса кластеризации клиентской базы разработан программный комплекс «Clusterization of the client base».

Назначение программного комплекса - проведения кластерного анализа для разделения клиентской базы на группы, в которых клиенты имеют схожие характеристики (схожую поведенческую историю). Формирование таких групп позволит оптимизировать рекламные кампании, предлагать соответствующие сопутствующие товары, а также оценивать эффективность принятия управленческих решений.

Среда разработки программного комплекса - PyCharm, язык программирования Python (версия 3.9).

В качестве базового алгоритма кластеризации был выбран метод к-средних. Выбор метода обусловлен рядом причин:

основное преимущество метода заключается в том, что он обладает линейной вычислительной сложностью, в отличие от алгоритма иерархической кластеризации и DBSCAN, что делает его более эффективным при обработке большого массива клиентских записей;

кластеры, формируемые методом к-средними, имеют четко определенные центроиды, что упрощает описание каждой группы контрагентов;

метод не требует определения дополнительных внешних параметров (кроме числа кластеров), например, таких как ε -окрестность и минимальное количество точек в окрестности для признания точки ключевой как в методе DBSCAN.

Таким образом выбор метода к-средних обосновывается его высокой вычислительной эффективностью, что критически важно для работы с большими базами данных маркетплейсов, и наглядностью в интерпретации на основе формируемых в методе центроидов.

Программный комплекс состоит из набора модулей каждый из которых предназначен для выполнения определенного этапа кластеризации.

1. «Модуль ввода и предобработки данных». Модуль предназначен для группировки и агрегации первичных данных, в случае, если этот этап необходим для дальнейшего анализа данных. Так, например, если выгрузка содержит информацию об отдельных операциях клиентов за определенный период, то целесообразно провести группировку данных по уникальному коду клиента и рассчитать итоговые метрики в каждой группе с целью формирования интегральных показателей. Также функционал модуля включает в себя заполнение пропусков (NaN). Данный этап является критически важным для предварительной обработки данных, поскольку пропущенные значения препятствуют проведению дальнейшего анализа. Стратегия заполнения пропусков определяется на основе типа измерительной шкалы переменной: если шкала категориальная (определяется как имеющая менее 10 уникальных значений [10]), пропуски заполняются модой; в случае интервальной шкалы для заполнения используется метод линейной интерполяции. При необходимости возможна стандартизация переменных, которая позволит привести все переменные к единому масштабу.

2. «Модуль определения оптимального числа кластеров». Данный этап является важным шагом в процедуре кластеризации, так как позволяет определить оптимальное число групп, наилучшим образом отражающих внутреннюю структуру данных. Для решения этой задачи в программном комплексе используется метод локтя. Согласно этому

методу, процедура кластеризации с использованием алгоритма k-средних повторяется несколько раз. На каждой итерации рассчитывается суммарная внутрикластерная дисперсия (WCSS), далее строится график зависимости WCSS от количества кластеров. Анализируя полученный график, аналитик определяет точку, в которой снижение WCSS замедляется. Количество кластеров, которому соответствует эта точка принимается за итоговый кластерный идентификатор.

3. «Модуль кластеризации». Проведение кластерного анализа с использованием алгоритма k-средних при выбранном числе кластеров и формирование нового столбца в исходном наборе данных, который будет содержать метку, показывающую принадлежность клиента к тому или иному кластеру.

4. Модуль оценки качества кластеризации. Данный модуль позволяет решить три ключевые задачи: оценка сплоченности объектов внутри кластеров; выявление наиболее информативных переменных; установление статистически значимых различий между кластерными центрами. С этой целью в модуле реализованы силуэтный метод и дисперсионный анализ.

Для оценки сплоченности объектов в кластере рассчитываются коэффициенты силуэта, которые позволяют аналитику сделать вывод о том, насколько хорошо каждый объект «соответствует» своему кластеру по сравнению с соседними, тем самым оценивая внутреннюю компактность групп.

В свою очередь, дисперсионный анализ (ANOVA) применяется для выявления наиболее информативных переменных. Проведение ANOVA позволяет определить, существует ли значимая разница между средними значениями каждой переменной в центрах различных кластеров, что, в конечном счете, помогает понять, какие именно характеристики лучше всего отличают одну клиентскую группу от другой.

Аналитик имеет возможность выполнять кластерный анализ многократно, последовательно исключая из анализа малоинформативные переменные с целью получения значимого и хорошо интерпретируемого результата.

После формирования окончательного варианта решения выполняется экспорт итогового датасета в формате *.csv. Полученный файл включает идентификаторы клиентов и соответствующие номера кластеров, к которым относятся клиенты, что позволяет эффективно интегрировать результаты анализа в системы CRM или BI-платформы. Таким образом, обеспечивается удобство дальнейшей обработки и использования полученной информации для выработки адресных маркетинговых кампаний, улучшения продуктовых предложений и оптимизации процессов взаимодействия с клиентами.

Для визуализации результата могут использоваться специализированные BI-системы, в частности Yandex DataLens (или аналоги). Использование такой платформы позволит создавать интерактивные дашборды, отображающие профили сформированных кластеров, ключевые различия между ними (например, с помощью столбчатых диаграмм или тепловых карт), а также динамически фильтровать данные, обеспечивая глубокое погружение в структуру клиентской базы для принятия дальнейших стратегических решений.

Для оценки масштабируемости программного комплекса было проведено тестирование его производительности при кластеризации клиентской базы с фиксированным количеством кластеров ($k=3$) и переменным объемом входных данных ($N_1=22\ 000$, $N_2=44\ 000$, $N_3=88\ 000$, $N_4=176\ 000$, $N_5=700\ 000$)

Результаты тестирования представлены в табл. 2. Они демонстрируют высокую вычислительную эффективность, что является ключевым фактором, позволяющим использовать программный комплекс для регулярного анализа постоянно растущей клиентской базы.

Таблица 2. Оценка производительности

| N | Время на проведения кластеризации (сек) |
|---------|---|
| 22 000 | 0,193 |
| 44 000 | 0,196 |
| 88 000 | 0,198 |
| 176 000 | 0,199 |
| 700 000 | 0,201 |

Результаты и обсуждения

Рассмотрим возможности программного комплекса при кластеризации клиентской базы маркетплейса по продаже спортивных товаров четырех брендов. Датафрейм содержит 21 894 записей и более 30 показателей, позволяющих сформировать детальный профиль каждого клиента.

Поля включают в себя три типа данных:

личные данные клиента, например, номер клиента, возраст, пол;

финансовые и поведенческие метрики, например, общая сумма выкупа, процент выкупа, общее количество совершенных операций, количество операций за период, сумма выкупа за период;

продуктовые предпочтения: предпочитаемые бренды, типы товаров.

Такой богатый набор данных позволяет программному комплексу формировать кластеры, отражающие реальное поведение и предпочтения покупателей спортивных товаров.

Датафрейм выгружается в программный комплекс и активизируется работа «Модуля ввода и предобработки данных», далее происходит агрегация информации по каждому клиенту за рассматриваемый период (в программе предусмотрен выбор полей для группировки, агрегации, а также выбор агрегирующих функций), заполнение пропущенных данных, отбор переменных для кластерного анализа, стандартизация (рис. 1). В результате был сформирован датафрейм, включающий 7 469 записей.

На следующем этапе строится график зависимости WCSS от количества кластеров (рис. 2). В примере в качестве оптимального

количества кластеров было выбрано три кластера.

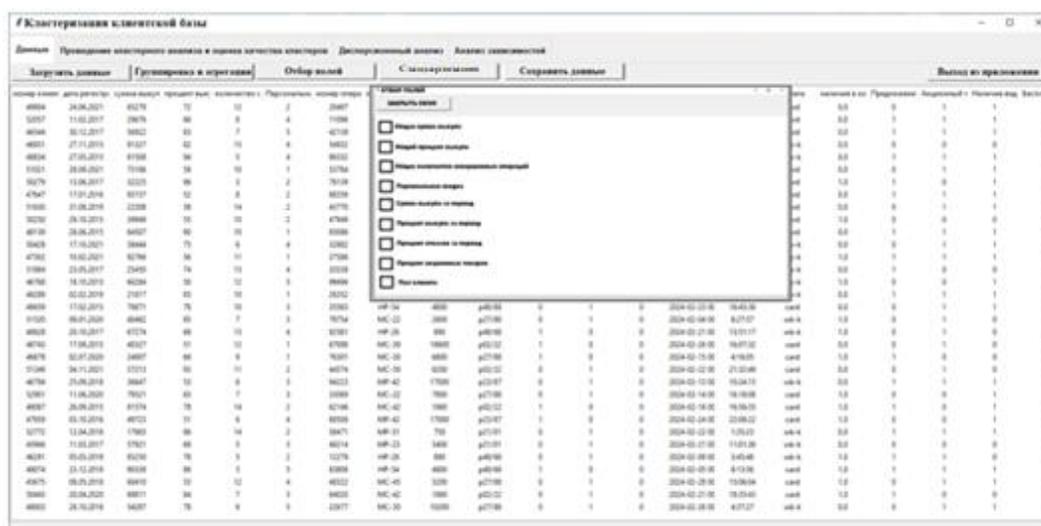


Рис. 1. Стартовое окно приложения

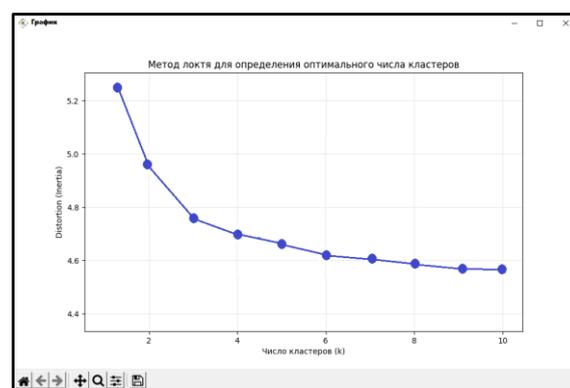


Рис. 2. Метод локтя

Результатом проведения процедуры кластерного анализа является датафрейм, который отличается от входного тем, что в нем сформирован новый столбец «Кластер», содержащий номер группы к которой относится каждый клиент. Пользователь имеет возможность выгрузить данный датафрейм в формате *.csv для дальнейшей обработки и визуализации результатов.

Для оценки качества кластеризации используется соответствующий модуль. Программа позволяет рассчитать силуэтные коэффициенты и провести дисперсионный анализ. В рассматриваемом примере итоговая таблица с силуэтными коэффициентами кластеров будет иметь вид (табл. 3).

Таблица 3. Силуэтные коэффициенты кластеров (значения выше 0,5 отмечены *)

| Характеристика | Кластер 1 | Кластер 2 | Кластер 3 |
|---|--------------------|--------------------|------------------------------|
| Размер | 2187 (30%) | 4177 (55%) | 1105 (15%) |
| Средний силуэтный коэффициент | 0,82* | 0,61* | 0,31 |
| Количество объектов с отрицательным коэффициентом | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) |
| Рекомендации | Выраженный кластер | Выраженный кластер | Это кластер требует внимания |

На основании данных таблицы можно сделать следующие выводы:

«Кластер 1» характеризуется наибольшими значениями среднего силуэтного коэффициента (0,82). Объекты внутри этого кластера демонстрируют высокую степень схожести друг с другом и низкую вероятность попадания в другие кластеры. Этот кластер является наиболее однородным среди всех выделенных групп.

«Кластер 2» характеризуется средним значением показателя силуэтного коэффициента, который составляет 0,61, что свидетельствует о достаточно хорошем уровне внутренней согласованности элементов.

«Кластер 3» отличается наименьшим средним значением силуэтного коэффициента (0,31), что говорит о низкой уверенности отнесения объектов именно к этому кластеру. Отсутствие объектов с отрицательными коэффициентами означает отсутствие явных ошибок классификации, однако низкое среднее значение сигнализирует о необходимости дополнительного изучения характеристик данного кластера.

Таким образом, большинство объектов уверенно распределены по соответствующим кластерам, за исключением третьего кластера, который нуждается в дополнительном внимании для повышения точности разделения объектов.

Исследователь имеет возможность оценивать статистическую значимость различий кластерных центров, сформировав таблицу с

результатами дисперсионного анализа (табл. 4).

Таблица 4. Фрагмент таблицы с результатами дисперсионного анализа

| Переменная | Межгрупповая дисперсия | Внутригрупповая дисперсия | F-статистика | P-value |
|--------------------------|------------------------|---------------------------|--------------|---------|
| Возраст | 3,56 | 0,99 | 3,56 | 0,006 |
| Пол | 1,66 | 0,99 | 1,66 | 0,156 |
| Общая сумма выкупа | 610,07 | 0,67 | 905,52 | 0,000 |
| Сумма выкупа за период | 723,11 | 0,61 | 1179,35 | 0,000 |
| Процент выкупа за период | 945,80 | 0,49 | 1915,33 | 0,000 |

Значения столбца p-value позволит определить неинформативные переменные ($p\text{-value} > 0,05$), которые исключаются из анализа и процедура кластеризации выполняется вновь. В таблице 3 примером неинформативной переменной является переменная пол.

По результатам кластеризации сформирован *.csv файл где каждому клиенту поставлен в соответствие номер кластера, к которому он относится. Полученный файл может быть выгружен в BI-систему, где формируется дашборд, содержащий визуализацию основных характеристик кластеров.

На рис. 3 изображен дашборд, созданный в Yandex DataLens, который содержит визуализацию основных метрик кластеров, в частности активна вкладка, содержащая информацию о первом кластере.

По данным дашборда можно дать следующую характеристику клиентам первого кластера:

- объем кластера 2 187 человек (30 % от общего числа покупателей);
- в кластер входят молодые покупатели (средний возраст 29 лет), соотношение мужчин и женщин одинаковое (52% - мужчины, 48% - женщины);
- средняя стоимость выкупленных заказов составляет 5 000 рублей, достаточно высокий процент выкупа товаров 75%;

- отдают предпочтение молодежным брендам с демократичными ценами (Бренд А и В);
- часто пользуются акциями при покупке товаров (90% акционных товаров);
- предпочитают стильную повседневную одежду, аксессуары и обувь для активного образа жизни.

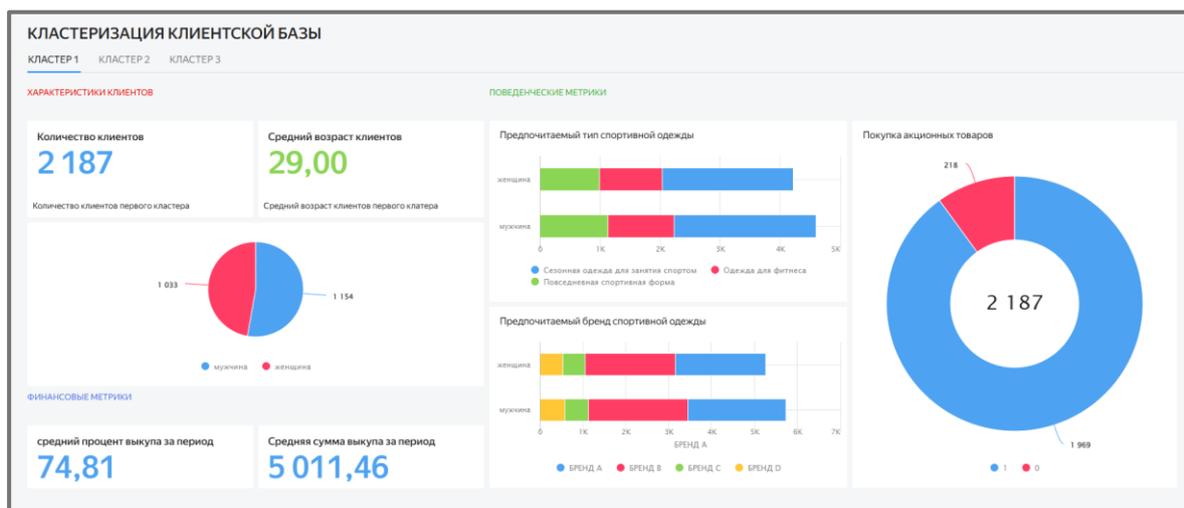


Рис. 3. Визуализация метрик первого кластера

Возможные рекомендации: клиентов этого кластера необходимо информировать о всех акциях и распродажах, а также предлагать скидки при покупке комплектов одежды и обуви из предпочитаемых брендов для увеличения среднего чека. Необходимы регулярные анонсы новых коллекций повседневной спортивной одежды и обуви для зала и города предпочитаемых брендов. Учитывая молодой возраст аудитории необходима более активная коммуникация и использование каналов, в которых представлена эта аудитория (социальные сети, мессенджеры, электронные сообщения).

По данным вкладок «Кластер 2» и «Кластер 3» рис. 3 можно произвести описание других кластеров и сформировать рекомендации для работы с клиентами, входящими в них.

Заключение

Использование кластерного анализа даёт аналитику возможность

получить представление о внутренних структурах и паттернах поведения клиентов, а также сформировать персонифицированную стратегию взаимодействия с каждым кластером.

Современные методы анализа данных и специализированное программное обеспечение поможет бизнесу повысить эффективность коммуникаций с клиентами, улучшив показатели лояльности и прибыльности бизнеса.

В настоящей статье описаны возможности программного комплекса, реализующего полную процедуру кластеризации клиентской базы. Продемонстрированы его возможности при кластеризации клиентской аудитории маркетплейса. Программный комплекс показал эффективность кластерного подхода для последующей оптимизации логистических и маркетинговых стратегий.

В перспективе программный комплекс может быть дополнен модулями выбора алгоритма кластеризации в зависимости от набора анализируемых данных.

Список литературы

1. Нигай Е.А. Управление «пожизненной ценностью клиента» (Lifetime Value) в деятельности бизнеса: методы оценки и стратегические аспекты принятия решений. Территория новых возможностей // Вестник Владивостокского государственного университета. 2024. Т. 16. № 3 (71). С. 54-65.

2. Хлебинский Н.Ю. Как системно работать с клиентскими данными и управлять удержанием клиентов, повышая показатели LTV и CRR // Интернет-маркетинг. 2021. № 1. С. 2-11.

3. Галямов А.Ф., Тархов С.В. Управление взаимодействием с клиентами коммерческой организации на основе методов сегментации и кластеризации клиентской базы // Вестник Уфимского государственного авиационного технического университета. 2014. Т. 18. № 4 (65). С. 149-156.

4. ABC и XYZ анализ продаж в 1С – автоматизация анализа продаж и управления продажами [Электронный ресурс] – URL: <https://www.it-sochi.ru/articles/kak-povysit-effektivnost-s-1s-abc-i-xyz-analiz-prodazh/> (дата обращения 21.01.2026).

5. Ярошенко Е.В. Особенности формирования клиентских сегментов для сетевых организаций в эпоху SMART // Открытое образование. 2017. Т. 21. № 1. С. 74-80.

6. Крюкова А.А., Пальмов С.В. Исследование применимости методов технологии DATA MINING для анализа клиентской базы телекоммуникационной компании //

Прикладная информатика. 2019. Т. 14. № 1 (79). С. 17-28.

7. Карпов Е.С. Сегментации клиентской базы методом k-средних основанном на построении графа k-ближайших соседей // World science: problems and innovations. Сборник статей XXXVI Международной научно-практической конференции. 2019. С. 30-32.

8. Петрова Н.С. Анализ данных клиентской базы для прогнозирования и оптимизации маркетинговых стратегий // Russian Economic Bulletin. 2024. Том 7. № 3. С. 148 – 154.

9. Трусевич И.В. Консолидация данных и интеллектуальный анализ данных клиентской базы корпоративных клиентов в банковском секторе экономики // Молодежь для науки и экономики: разработки и перспективы. Сборник научных статей VI международного форума молодых ученых. под науч. ред. А. П. Бобовича. 2017. С. 99-103.

10. Фомина Е.Е. Факторный анализ и категориальный метод главных компонент: сравнительный анализ и практическое применение для обработки результатов анкетирования // Гуманитарный вестник. 2017. № 10 (60). С. 3.

References

1. Nigaj E.A. Upravlenie «pozhitzennoj cennost`yu klienta» (Lifetime Value) v deyatel`nosti biznesa: metody` ocenki i strategicheskie aspekty` prinyatiya reshenij. Territoriya novy`x vozmozhnostej // Vestnik Vladivostokskogo gosudarstvennogo universiteta. 2024. T. 16. № 3 (71). S. 54-65.

2. Xlebinskij N.Yu. Kak sistemno rabotat` s klientskimi dannymi i upravlyat` uderzhanie klientov, povu`shaya pokazateli LTV i CRR // Internet-marketing. 2021. № 1. S. 2-11.

3. Galyamov A.F., Tarxov S.V. Upravlenie vzaimodejstviem s klientami kommercheskoj organizacii na osnove metodov segmentacii i klasterizacii klientskoj bazy` // Vestnik Ufimskogo gosudarstvennogo aviacionnogo texnicheskogo universiteta. 2014. T. 18. № 4 (65). S. 149-156.

4. AVS i XYZ analiz prodazh v 1S — avtomatizaciya analiza prodazh i upravleniya prodazhami [E`lektronny`j resurs] – URL: <https://www.it-sochi.ru/articles/kak-povysit-effektivnost-s-1s-abc-i-xyz-analiz-prodazh/> (data obrashheniya 24.08.2025).

5. Yaroshenko E.V. Osobennosti formirovaniya klientskix segmentov dlya setevy`x organizacij v e`poxu SMART // Otkry`toe obrazovanie. 2017. T. 21. № 1. S. 74-80.

6. Kryukova A.A., Pal`mov S.V. Issledovanie primenimosti metodov texnologii DATA MINING dlya analiza klientskoj bazy` telekommunikacionnoj kompanii // Prikladnaya informatika. 2019. T. 14. № 1 (79). S. 17-28.

7. Karpov E.S. Segmentacii klientskoj bazy` metodom k-srednix osnovannom na postroenii grafa k-blizhajshix sosedej // World science: problems and innovations. Sbornik statej XXXVI Mezhdunarodnoj nauchno-prakticheskoy konferencii. 2019. S. 30-32.

8. Petrova N.S. Analiz dannyx klientskoj bazy` dlya prognozirovaniya i optimizacii marketingovy`x strategij // Russian Economic Bulletin. 2024. Tom 7. № 3. S. 148 – 154.

9. Trusevich I.V. Konsolidaciya dannyx i intellektual`ny`j analiz dannyx klientskoj bazy` korporativny`x klientov v bankovskom sektore e`konomiki // Molodezh` dlya nauki i e`konomiki: razrabotki i perspektivy`. Sbornik nauchny`x statej VI mezhdunarodnogo foruma molody`x ucheny`x. pod nauch. red. A. P. Bobovicha. 2017. S. 99-103.

10. Fomina E.E. Faktorny`j analiz i kategorial`ny`j metod glavny`x komponent: sravnitel`ny`j analiz i prakticheskoe primenenie dlya obrabotki rezul`tatov anketirovaniya // Gumanitarny`j vestnik. 2017. № 10 (60). S. 3.