

ТЕХНИЧЕСКИЕ НАУКИ / TECHNICAL SCIENCES

DOI: [10.15507/2311-2468.013.202501.053-063](https://doi.org/10.15507/2311-2468.013.202501.053-063)

eISSN 2311-2468

EDN: <https://elibrary.ru/vzpbva><https://ogarev-online.ru>

УДК / UDC 69.05

Оригинальная статья / Original article

ИСПОЛЬЗОВАНИЕ ПЛАТФОРМЫ ВИЗУАЛЬНОЙ РАЗРАБОТКИ СЦЕНАРИЕВ ДЛЯ МОДЕЛИРОВАНИЯ ПРОЦЕССОВ РЫНКА НЕДВИЖИМОСТИ

Д. К. Егорова, Р. В. Денисов ✉Национальный исследовательский Мордовский государственный университет,
Саранск, Россия✉ ar4yp@yandex.ru

Аннотация

Введение. Рынок недвижимости – ключевой сектор экономики с высокой динамикой цен, зависимостью от макроэкономических факторов и сложностью прогнозирования. Традиционные методы анализа требуют много времени и ресурсов, что ограничивает их применение. Использование low-code платформ позволяет сократить затраты на разработку моделей и сделать инструменты анализа доступными для специалистов без углубленных навыков программирования. Цель исследования – продемонстрировать использование KNIME для прогнозирования стоимости объектов недвижимости и их классификации; оценить, насколько точны модели и насколько они полезны на практике.

Материалы и методы. Реализовано моделирование процессов рынка недвижимости с помощью платформы визуальной разработки сценариев KNIME Analytics Platform. Данные об объектах недвижимости собраны средствами Python библиотеки Cianparser, в KNIME Analytics Platform реализованы методы регрессионного анализа и визуализации данных.

Результаты исследования. Построены линейная и полиномиальная регрессия цен на недвижимость по заданным параметрам, выполнена кластеризация объектов недвижимости и визуализация полученных результатов. Кластеризация выявила три группы объектов, коррелирующих с локацией и инфраструктурой.

Обсуждение и заключение. KNIME подтвердил эффективность как low-code инструмент для анализа рынка недвижимости. Материалы статьи могут быть полезны для понимания динамики рынка недвижимости и прогнозирования его будущих тенденций.

Ключевые слова: KNIME Analytics Platform, регрессия, методы снижения размерности, метод главных компонент, кластеризация

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования: Егорова Д. К., Денисов Р. В. Использование платформы визуальной разработки сценариев для моделирования процессов рынка недвижимости // Огарёв-online. 2025. Т. 13, No 1. С. 53–63. <https://doi.org/10.15507/2311-2468.013.202501.053-063>

© Егорова Д. К., Денисов Р. В., 2025

Контент доступен по лицензии Creative Commons Attribution 4.0 License.
This work is licensed under a Creative Commons Attribution 4.0 License.

VISUAL SCENARIO DEVELOPMENT PLATFORM USED TO MODEL REAL ESTATE MARKET PROCESSES

D. K. Egorova, R. V. Denisov

National Research Mordovia State University,

Saransk, Russia

✉ ar4yp@yandex.ru

Abstract

Introduction. The real estate market is a key sector of the economy with high price dynamics, dependence on macroeconomic factors and complexity of forecasting. Traditional analysis methods require a lot of time and resources, which limits their application. Using low-code platforms allows you to reduce the cost of developing models and make analysis tools accessible to specialists without advanced programming skills. The purpose of the study is to demonstrate the use of KNIME to predict the value of real estate and their classification. In addition, it is necessary to assess how accurate the models are and how useful they are in practice.

Materials and Methods. Real estate market processes were modeled using KNIME Analytics Platform for visual scenario development. Real estate data is collected using the Cyanparser Python library, and regression analysis and data visualization methods are implemented in KNIME Analytics Platform.

Results. Linear and polynomial regression of real estate prices according to specified parameters is constructed, clusterization of real estate objects and visualization of the results are performed. Clustering revealed three groups of objects correlating with location and infrastructure.

Discussion and Conclusion. KNIME has confirmed its effectiveness as a low-code tool for analyzing the real estate market. The materials of the article can be useful for understanding the dynamics of the real estate market and forecasting its future trends.

Keywords: KNIME Analytics Platform, regression, dimensionality reduction methods, principal component analysis, clustering

Conflict of interest. The authors declare no conflict of interest.

For citation: Egorova D.K., Denisov R.V. Visual Scenario Development Platform Used to Model Real Estate Market Processes. *Ogarev-online*. 2025;13(1):53–63. <https://doi.org/10.15507/2311-2468.013.202501.053-063>

ВВЕДЕНИЕ

В настоящее время анализ данных и моделирование процессов играют ключевую роль в принятии эффективных решений. Рынок недвижимости не является исключением. Его сложность и высокая волатильность требуют применения определенных инструментов и методов для понимания динамики и прогнозирования будущих тенденций.

Весьма актуальным здесь является использование инструментария low-code, позволяющего сократить время разработки моделей для специалистов без глубоких навыков программирования. В частности, «процесс анализа следует “собрать” из своего рода узлов-nodes, в которых используются библиотеки Python, R, JavaScript» [1].

Примером такого инструмента является аналитическая платформа визуальной разработки сценариев KNIME¹, которая может использоваться для моделирования процессов рынка недвижимости – извлечения данных, построения модели прогнозирования цен на

¹ KNIME Analytics Platform [Электронный ресурс]. URL: <https://www.knime.com> (дата обращения: 01.02.2025).

недвижимость, кластеризации объектов недвижимости по нескольким параметрам, визуализации результатов.

Цель исследования – продемонстрировать применение KNIME для прогнозирования цен и кластеризации объектов недвижимости, оценив точность моделей и их практическое использование.

ОБЗОР ЛИТЕРАТУРЫ

В данном направлении проведено множество исследований. Например, предложен подход к определению стоимости объектов недвижимости на основе метода бинарного кодирования качественных характеристик объекта недвижимости, выполнен кластерный анализ определения схожих по характеристикам объектов, построена регрессионная модель стоимости объекта недвижимости [2]. В. Н. Деркаченко построена модель краткосрочного прогноза средней цены одного квадратного метра жилья [3].

Для анализа рынка недвижимости М. А. Зуев использовал метод кластеризации k-means, чтобы найти оптимальное число кластеров в рамках решения задачи для одного мегаполиса [4]. На основе исследования эластичности цен на жилье по доходам и типизации рынков Г. М. Стерник построил линейные модели прогнозирования цен на жилье при наличии прогнозов динамики душевых доходов населения [5].

В работах² [6; 7] приведен обзор и сравнительный анализ инструментария Data Mining, таких как Weka, RapidMiner, TANAGRA, Orange, Deductor, WizWhy KNIME и т. д. Рассмотрены принципы работы данных инструментов, представлены основные критерии для их сравнения.

Средствами KNIME реализована кластеризация однородных районов в рамках крупных территориальных образований, которая может быть использована при организации обоснованного государственного финансирования развития территорий [8].

При этом потенциал KNIME для задач кластеризации и прогнозирования в контексте региональных рынков недвижимости изучен недостаточно, что определяет научную новизну исследования.

МАТЕРИАЛЫ И МЕТОДЫ

В данном исследовании для моделирования процессов на рынке недвижимости использовалась аналитическая платформа визуальной разработки сценариев KNIME.

Для анализа рынка недвижимости в г. Саранске использованы данные, собранные с крупнейшего в России ресурса объявлений о продаже и аренде жилой, загородной и коммерческой недвижимости – сервиса ЦИАН. Сбор данных осуществлялся средствами Python библиотеки Cianparser³. В результате получены данные о 1 530 объектах недвижимости, включая квартиры, которые были выставлены на продажу на платформе по состоянию на 1 июня 2024 г. Признаки объектов недвижимости, используемые в анализе, представлены на рисунке 1.

floor Num_	floor... Numbe...	room... Number...	total... Number...	price Number (i...	district String	street String	hou... String	latit... String	longitude String
4	6	1	28	2300000	Ленинский	Попова	74Б	54.1732...	45.12248
3	5	1	32.5	2400000	Пролетар...	Семашко	9	54.2288...	45.12918
5	10	1	39.9	4650000	Ленинский	Кирова	31	54.1740...	45.17840
3	5	1	27.7	2550000	Ленинский	Республика...	48	54.1722...	45.17531
5	10	1	46.5	6550000	Ленинский	Московская	34	54.1760...	45.18656
8	10	1	39.1	4990000	Ленинский	Гагарина	92	54.1770...	45.15365

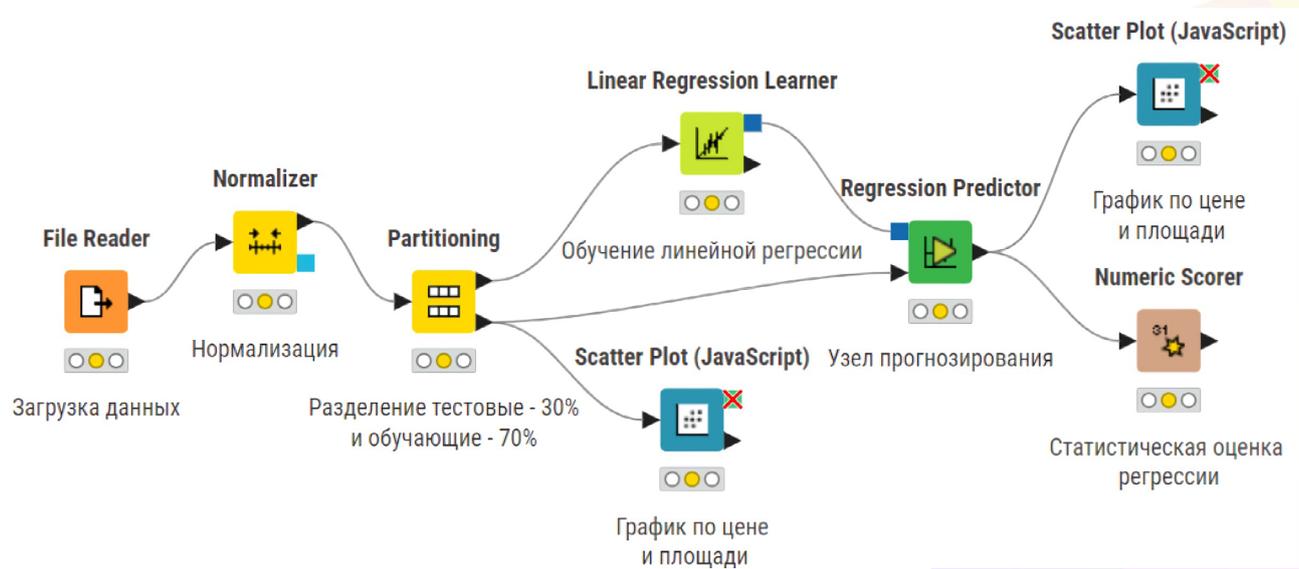
Р и с. 1. Признаки, получаемые в ходе сбора данных
F i g. 1. Features obtained during data collection

Источник: составлен авторами на основе данных сервиса ЦИАН.
Sources: compiled by the authors based on data from the CIAN service.

Для определения географических координат объектов недвижимости использован API сайта openstreetmap.org⁴. Из 1 530 собранных объявлений координаты удалось извлечь для 1 455 объектов.

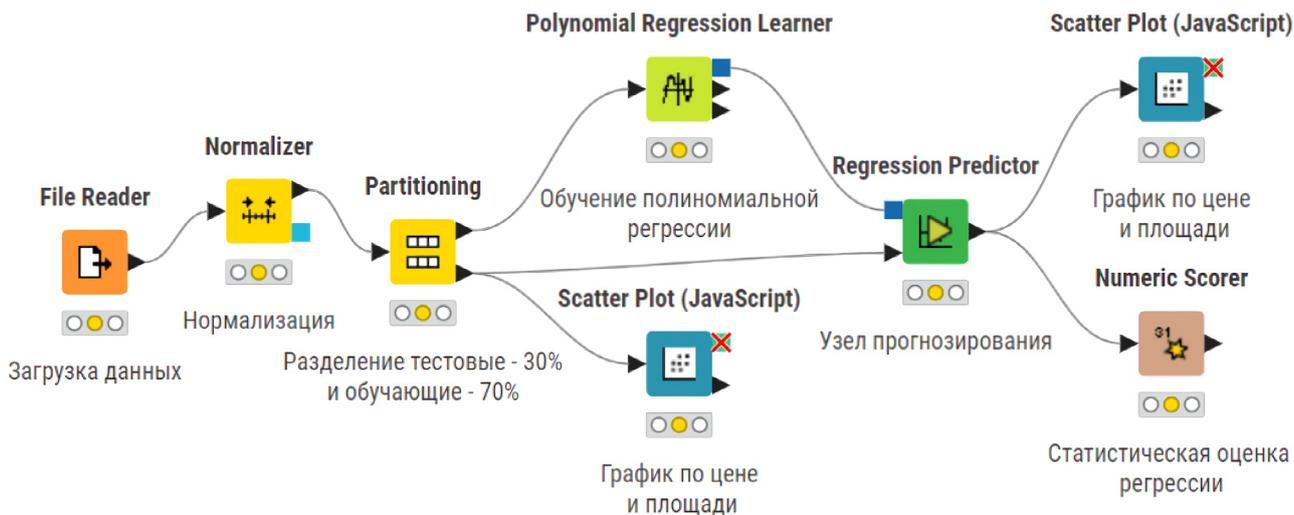
Моделирование в KNIME Analytics Platform. На рисунках 2, 3 представлены рабочие процессы KNIME, осуществляющие прогноз цен на недвижимость по заданным параметрам с помощью линейной и полиномиальной регрессии.

Независимыми переменными, в силу отсутствия попарной корреляции и большей практической значимости, здесь выбраны этаж, общее количество этажей здания, количество комнат, площадь. Зависимой переменной – цена. На рисунках 3, 4 приведены оценки моделей, на основании которых можно сделать вывод об удовлетворительном результате моделирования.



Р и с. 2. Процесс моделирования линейной регрессии
F i g. 2. The linear regression modeling process

Источник: рисунки 2–7, 9–13 составлены авторами.
Sources: figures 2–7, 9–13 are compiled by the authors.



Р и с. 3. Процесс моделирования полиномиальной регрессии
 F i g. 3. Polynomial regression modeling process

#	RowID	Prediction (price) <i>Number (double)</i>
1	R^2	0.776
2	mean absolute error	0.046
3	mean squared error	0.005
4	root mean squared error	0.068

Р и с. 4. Оценки линейной регрессии
 F i g. 4. Linear regression estimates

#	RowID	Prediction (price) <i>Number (double)</i>
1	R^2	0.755
2	mean absolute error	0.047
3	mean squared error	0.005
4	root mean squared error	0.069

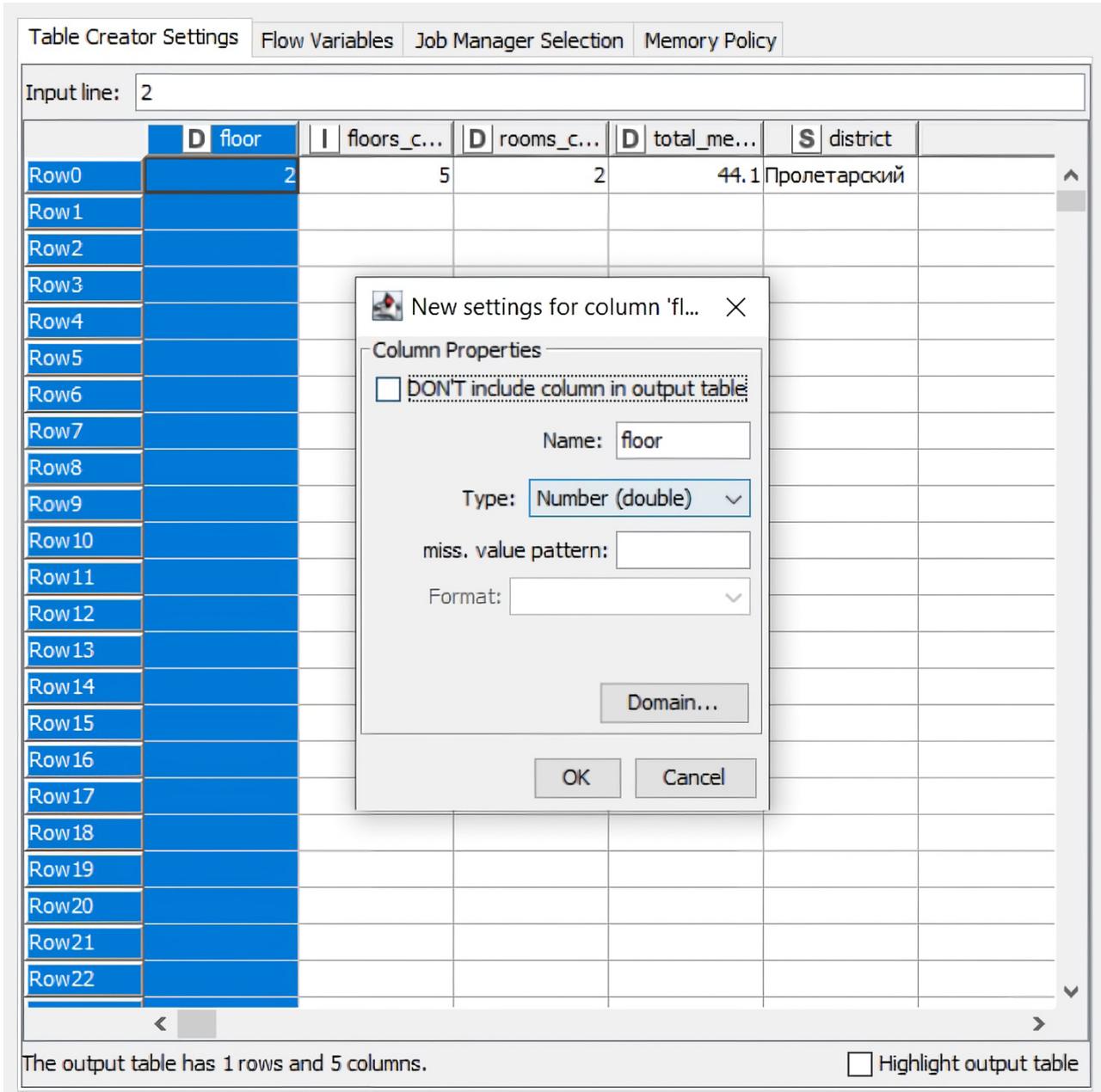
Р и с. 5. Оценки полимерной регрессии
 F i g. 5. Polynomial regression estimates

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

В рабочий процесс KNIME импортирован узел Table Creator, который позволяет вводить параметры объекта (независимые переменные) и получать прогноз, чтобы использовать полученные модели для прогнозирования цен. Конфигурация узла содержит столбцы, тип данных, названия которых совпадают с параметрами объекта недвижимости (рис. 6).

Для тестирования работоспособности модели вводились данные реальных объектов, затем результаты сравнивались. Например, при вводе данных объекта, расположенного по адресу г. о. Саранск, ул. Веселовского, д. 26, цена по объявлению на сервисе ЦИАН составляла 3 500 000 руб., а прогнозируемая цена по полиномиальной регрессии 3 616 320 677 руб., что, с некоторой погрешностью, соответствует рыночным ожиданиям (рис. 7).

Кроме этого, сравнивался подобный функционал прогнозирования цен на недвижимость известных сервисов. Например, сервис «Яндекс.Недвижимость» возможности прогнозирования цен не представляет, а оценка вышеупомянутой квартиры на сервисе ЦИАН совпала с данными моделирования в KNIME (рис. 8).



Р и с. 6. Конфигурация узла Table Creator
F i g. 6. Table Creator node configuration

#	RowID	floor <i>Number (double)</i>	floors_count <i>Number (integer)</i>	rooms_count <i>Number (double)</i>	total_meters <i>Number (double)</i>	district <i>String</i>	Prediction (price) <i>Number (double)</i>
1	Row0	2	5	2	44.1	Пролетарский	3,616,320.877

Р и с. 7. Работа узла Table Creator
F i g. 7. Operation of the Table Creator node



Оценка этой квартиры от Циан

[Смотреть полный отчет](#)

Поможет покупателю поторговаться, а продавцу — продать квартиру выгоднее и быстрее

3,6 млн ₽

Оценка Циан ⓘ

3,5 млн ₽

Цена в объявлении

Хорошая цена

Изменение цены

Неделя Месяц Полгода

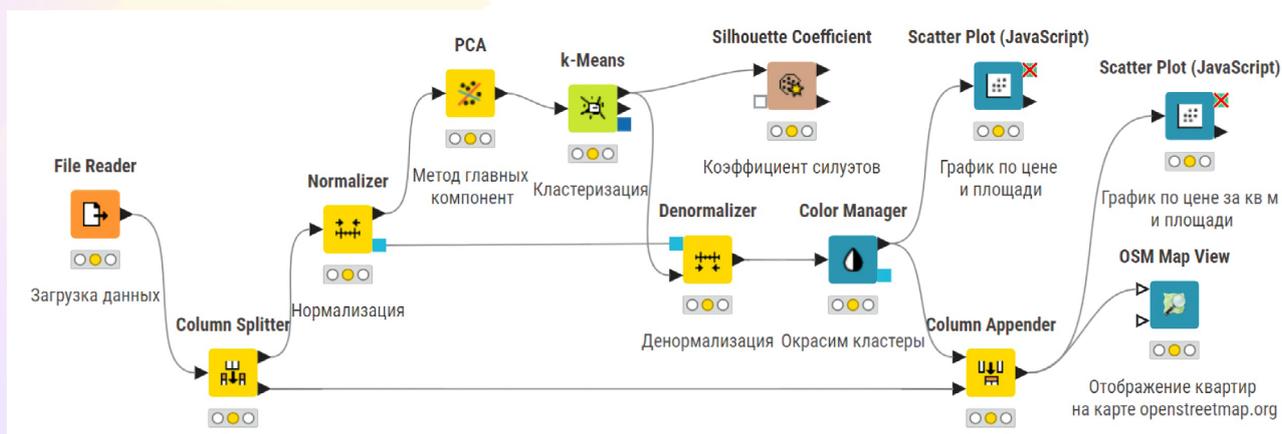


Р и с. 8. Сервис оценки невидимости ЦИАИ
 Fig. 8. CIAN invisibility assessment service

Источник: сервис ЦИАИ.
 Sources: CIAN service.

Методы снижения размерности и кластеризации данных. Визуализация данных также является весьма эффективной при анализе рынка недвижимости. Визуализацию можно выполнить последовательно, применив один из методов снижения размерности признакового пространства, а затем выполнив кластеризацию. Данные методы реализованы в KNIME.

Рабочий процесс визуализации данных недвижимости представлен на рисунке 9. После чтения и нормализации данных был применен метод главных компонент, а затем проведена кластеризация методом k -средних. Количество кластеров ($k = 3$) получено путем применения метода силуэтов (рис. 10).

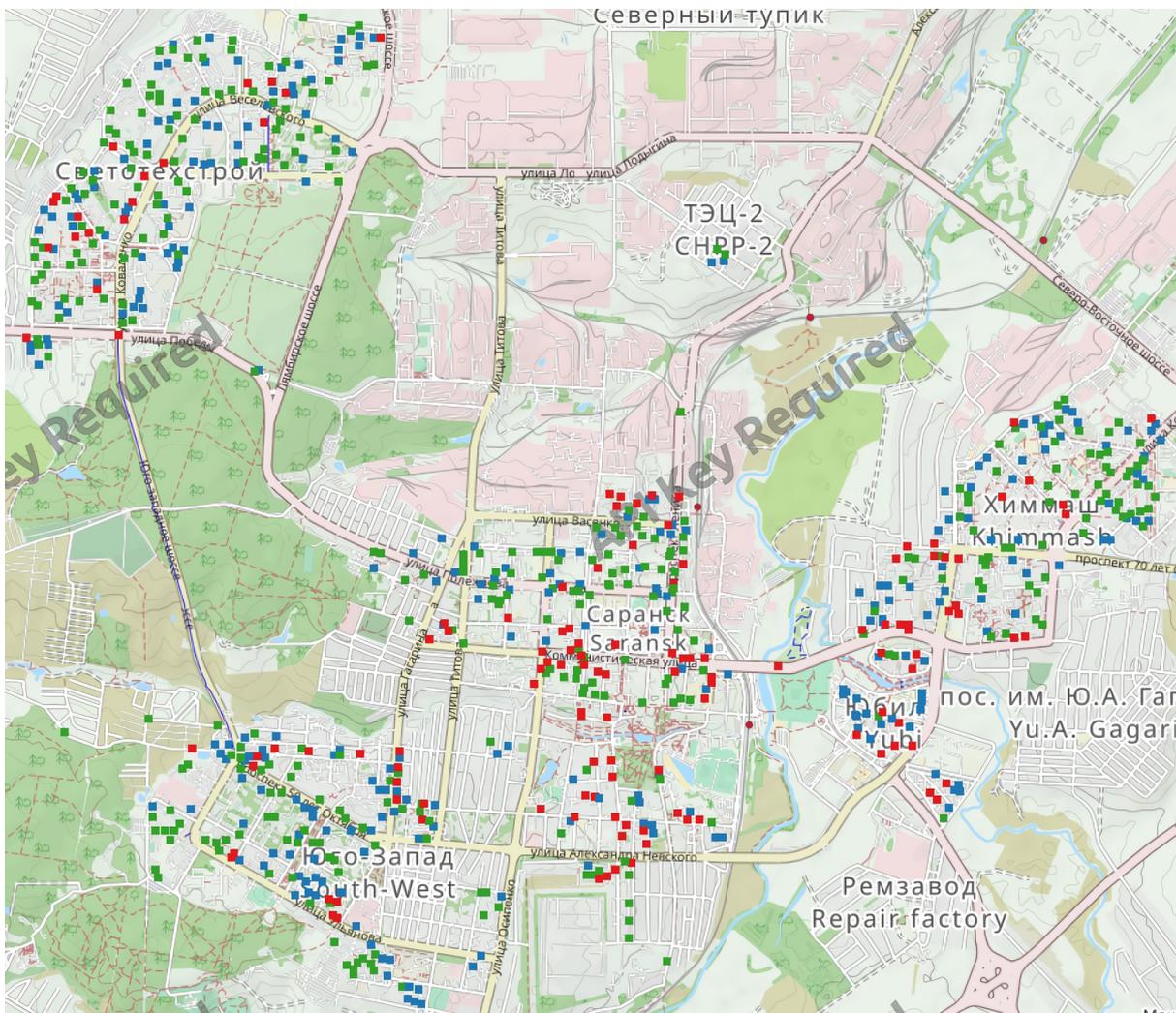


Р и с. 9. Рабочий процесс визуализации данных
 Fig. 9. Data visualization workflow

#	RowID	Mean Silhouette Coefficient <i>Number (double)</i>
1	cluster_2	0.397
2	cluster_1	0.189
3	cluster_0	0.412
4	Overall	0.367

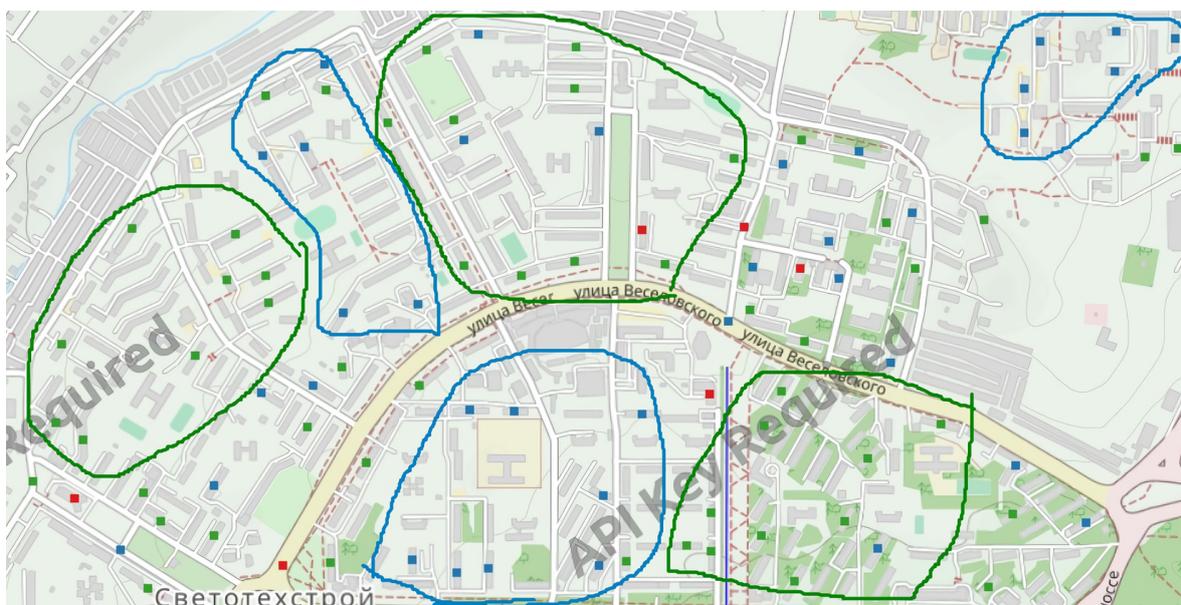
Р и с. 10. Оценка кластеризации методом силуэтов
 Fig. 10. Evaluation of clustering using the silhouette method

Визуализация осуществлялась в том числе с помощью узла OSM Map View, который предоставляет интерактивный доступ к картам OpenStreetMap (openstreetmap.org). Кластеры на карте openstreetmap окрашены разными цветами и представлены на рисунке 11.

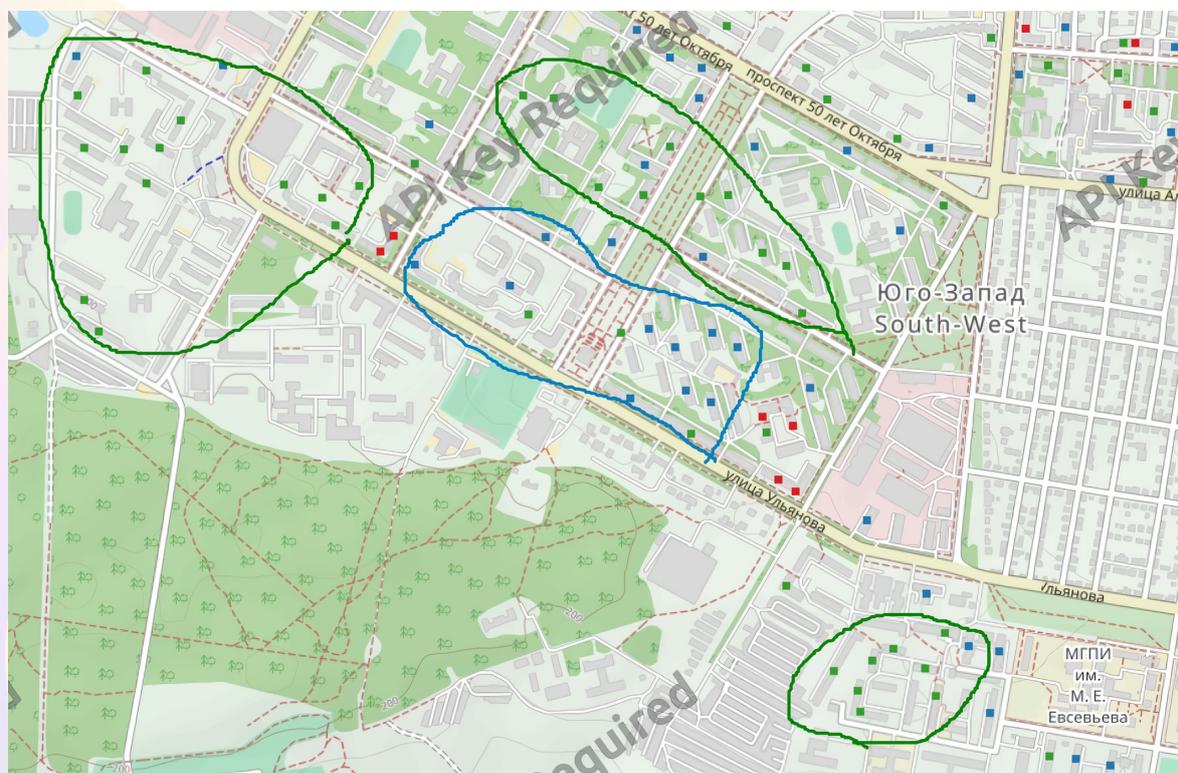


Р и с. 11. Кластеры на карте openstreetmap
 Fig. 11. Clusters on the openstreetmap

В ходе анализа результатов визуализации выявлено, что, например, квартиры в районах Светотехстрой и Юго-Запад чаще объединяются в один кластер в пределах одного жилого массива (несколько домов находящихся рядом, образующих «колодец», в непосредственной близости от школы, детского сада, магазина). Кластеры района Светотехстрой представлены на рисунке 12, кластеры района Юго-Запад – на рисунке 13. Результаты кластеризации коррелируют с визуальной оценкой интерьера квартир из разных кластеров.



Р и с. 12. Кластеры района Светотехстрой
F i g. 12. Clusters of the Svetotekhsstroy district



Р и с. 13. Кластеры района Юго-Запад
F i g. 13. Clusters of the South-West district

ОБСУЖДЕНИЕ И ЗАКЛЮЧЕНИЕ

Одним из основных преимуществ использования KNIME Analytics Platform является то, что это low-code инструмент, который позволяет пользователю минимизировать программный код, осуществляя построение моделей с помощью визуальных интерфейсов и готовых модулей. Именно low-code подход позволяет быстро создавать и изменять модели рынка недвижимости и может быть интересен специалистам, объектом исследования которых является рынок недвижимости.

В данной статье продемонстрировано применение KNIME Analytics Platform для моделирования процессов рынка недвижимости; построены рабочие процессы, реализующие линейную и полиномиальную модели регрессии прогноза цены объекта недвижимости по нескольким параметрам; проведено сравнение работы построенных моделей с данными сервисов ЦИАН и «Яндекс.Недвижимость»; выполнена визуализация кластерного анализа объектов недвижимости на карте OpenStreetMap.

Однако следует отметить, что проведенное исследование имеет определенные ограничения. Во-первых, не были учтены макроэкономические факторы, такие как инфляция и ставки кредитования, которые могут существенно влиять на рынок недвижимости. Во-вторых, данные были ограничены одним городом и конкретным периодом, что актуализирует необходимость дальнейших исследований для более полного и всестороннего анализа рынка недвижимости.

СПИСОК ЛИТЕРАТУРЫ

1. Егорова Д. К., Заварюхина Ю. В. Применение инструментария KNIME Analytics Platform для анализа соответствия рабочих программ учебных дисциплин требованиям работодателей // *Ogarev-online*. 2023. Т. 11, № 16. URL: <https://cyberleninka.ru/article/n/primeneniye-instrumentariya-knime-analytics-platform-dlya-analiza-sootvetstviya-rabochih-programm-uchebnyh-distiplin-trebovaniyam> (дата обращения: 20.02.2025).
2. Прогнозирование стоимости объектов недвижимости на основе комплексного анализа их свойств / О. В. Савина [и др.] // *Прикаспийский журнал: управление и высокие технологии*. 2019. № 4. С. 60–70. <https://doi.org/10.21672/2074-1707.2019.48.4.060-070>
3. Деркаченко В. Н. Прогнозирование и кластерный анализ развития регионального рынка жилой недвижимости // *Научно-методический электронный журнал «Концепт»*. 2014. Т. 20. С. 11–15. URL: <http://e-koncept.ru/2014/54262.htm> (дата обращения: 20.02.2025).
4. Зуев М. А., Шибяев В. М., Баланев К. С. Разработка модели K-Means для выявления наиболее выгодных предложений на рынке недвижимости Москвы // *Информатика. Экономика. Управление*. 2024. Т. 3, № 2. С. 212–218. <https://doi.org/10.47813/2782-5280-2024-3-2-0212-0218>
5. Стерник Г. М. Методика прогнозирования цен на жилье в зависимости от типа рынка // *Имущественные отношения в Российской Федерации*. 2011. № 1. С. 43–47. EDN: *NCCIRR*
6. Омарова Ш. Е., Медеубаева А. М. Сравнительный анализ инструментов Data Mining // *Заметки ученого*. 2020. № 11. С. 185–193. EDN *QYUTBN*
7. Пальмов С. В., Диязитдинова А. А., Артюшкина Е. С. Сравнительный анализ возможностей интеллектуальных систем при выявлении скрытых закономерностей в данных // *Электросвязь*. 2020. № 2. С. 52–58. <https://doi.org/10.34832/ELSV.2020.3.2.008>
8. Разработка подхода к кластеризации районов на базе инструментального средства машинного обучения KNIME / Е. М. Смирнова [и др.] // *Известия высших учебных заведений. Серия «Экономика, финансы и управление производством» [Ивэкофин]*. 2021. № 4. С. 165–175. URL: <https://ecofin-isuct.ru/article/view/4079> (дата обращения: 20.02.2025).

REFERENCES

1. Egorova D.K., Zavarukhina Y.V. Application of KNIME Analytics Platform Tools to Analyze the Compliance of Syllabuses with the Requirements of Employers. *Ogarev-online*. 2023;11(16). (In Russ.,

abstract in Eng.) Available at: <https://cyberleninka.ru/article/n/primeneniye-instrumentariya-knime-analytics-platform-dlya-analiza-sootvetstviya-rabochih-programm-uchebnyh-disttsiplin-trebovaniyam> (accessed 20.02.2025).

2. Savina O.V., Malikov V.P., Sadovnikova N. P., Parygin D.S., Mityagin S.A., Voronin D.Yu. Forecasting the Value of Real on the Basis of a Comprehensive Analysis of its Properties. *Caspian Journal: Control and High Technologies*. 2019;(4):60–70. (In Russ., abstract in Eng.) <https://doi.org/10.21672/2074-1707.2019.48.4.060-070>

3. Derkachenko V.N. Prognostication and the Cluster Analysis of the Development of the Regional Market for Habitable Real Estate. *Cientific and Methodological Electronic Journal "Koncept"*. 2014;20:11–15. (In Russ., abstract in Eng.) Available at: <http://e-koncept.ru/2014/54262.htm> (accessed 20.02.2025).

4. Zuev M.A., Shibaev V.M., Balanov K.S. Development of the K-Means Model to Identify the Most Profitable Offers on the Moscow Real Estate Market. *Informatics. Economics. Management*. 2024;3(2):212–218. (In Russ., abstract in Eng.) <https://doi.org/10.47813/2782-5280-2024-3-2-0212-0218>

5. Sternik G.M. Forecasting Techniques in Housing Prices Depending on the Type Market. *Property Relations in the Russian Federation*. 2011;(1):43–47. (In Russ., abstract in Eng.) EDN: NCCIRR

6. Omarova Sh.E., Medeubayeva A.M. Comparative Analysis of Data Mining Tools. *Zametki uchenogo*. 2020;(11):185–193 (In Russ., abstract in Eng.) EDN QYUTBN

7. Palmov S.V., Diyazitdinova A.A., Artyushkina E.S. Comparative Analysis of the Intelligent Systems Capabilities in Hidden Patterns Extracting. *Electrosvyaz*. 2020;(2):52–58. (In Russ., abstract in Eng.) <https://doi.org/10.34832/ELSV.2020.3.2.008>

8. Smirnova E.M., Valinurova A.A., Danilova S.V., Valinurov T.R. Development of an Approach to Clustering of Districts Based on the KNIME Machine Learning Tool. *Ivecofin*. 2021;(4):165–175. (In Russ., abstract in Eng.) Available at: <https://ecofin-isuct.ru/article/view/4079> (accessed 20.02.2025).

ИНФОРМАЦИЯ ОБ АВТОРЕ

Егорова Дарья Константиновна, кандидат физико-математических наук, доцент кафедры прикладной математики Национального исследовательского Мордовского государственного университета (430005, Российская Федерация, г. Саранск, ул. Большевикская, д. 68), ORCID: <https://orcid.org/0000-0002-3392-6761>, SPIN-код: 9427-8154, egorovadk@mail.ru

Денисов Роман Владимирович, магистрант факультета математики и информационных технологий Национального исследовательского Мордовского государственного университета (430005, Российская Федерация, г. Саранск, ул. Большевикская, д. 68), ORCID: <https://orcid.org/0009-0001-9652-0996>, ar4yp@yandex.ru

ЗАЯВЛЕННЫЙ ВКЛАД АВТОРОВ

Д. К. Егорова – разработка концепции; написание рукописи – рецензирование и редактирование.

Р. В. Денисов – проведение исследования; разработка программного обеспечения; написание черновика рукописи.

Все авторы прочитали и одобрили окончательный вариант рукописи.

Поступила 25.02.2025; одобрена после рецензирования 05.03.2025; принята к публикации 10.03.2025.

INFORMATION ABOUT THE AUTHOR

Daria K. Egorova, Cand.Sci. (Phys.-Math.), Associate Professor of the Department of Applied Mathematics, National Research Mordovia State University (68 Bolshevistskaya St., Saransk 430005, Russian Federation), ORCID: <https://orcid.org/0000-0002-3392-6761>, SPIN-code: 9427-8154, egorovadk@mail.ru

Roman V. Denisov, Master's Student of the Faculty of Mathematics and Information Technology, National Research Mordovia State University (68 Bolshevistskaya St., Saransk 430005, Russian Federation), ORCID: <https://orcid.org/0009-0001-9652-0996>, ar4yp@yandex.ru

CONTRIBUTION OF THE AUTHORS

D. K. Egorova – conceptualization; writing – review and editing.

R. V. Denisov – investigation, software; writing – original draft preparation.

The authors have read and approved the final manuscript.

Submitted 25.02.2025; revised 05.03.2025; accepted 10.03.2025.