PACIFIC RIM PROPERTY RESEARCH JOURNAL
2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327

Enhancing Condominium Rental Classification and Income Prediction through Optimized Modeling Techniques

Nashat Ali V Almasria[1], Tawfiq Tahtamoni[2], Hala Tawaha[3], Heba Hasan Alenezi[4]
[1]A'Sharqiyah University (ASU), [2]Oman, School of Global Business, Arcadia University, USA, [3]The University of New Orleans USA, [4]Independent Researcher- Data Analyst

[1]Email: nashat.almasria@asu.edu.om
[2]Email: Tahtamonit@arcadia.edu
[3]Email: hatawaha@uno.edu
[4]Email: hebaalhasan38@gmail.com

Abstract

Real estate is one of the economic pillars of countries. Especially the condominiums that governments allocate for investment through rental income. In this study, we propose a novel model to categorize and estimate condominium rental income (CRI) using innovative techniques, including the Z-score transform (ZT), correlation coefficients (C-C), fuzzy C-mean clustering (FCM), and linear Regression (LR). The proposed novel model employs a set of specific features for this purpose and posits an important hypothesis regarding the role of condominium market values (CMV), together with their volume measures, represented by the condominium gross square (CGS) and total units (TUN), in linearly estimating the correlated CRI. Through this research, we demonstrate the validity of our hypothesis by systematically performing LR on the selected features: we first cluster CGS, TUN, CMV, and CRI into distinct classes using FCM, then apply LR separately to each resulting class. Our research is guided by a practical study applied to the data recorded in the Data.gov Catalog Harvesting - New York State (DCH-NY) for the fiscal years 2008–2009, 2009–2010, 2010–2011, and 2011–2012 (Data.gov, 2025). We built a dataset comprising five parts; each part contains condominium data for one of five New York cities, considering only the best numerically correlated features with CRI. The system was applied and tested, achieving competitive accuracy in categorization and estimation. The proposed model achieves competitive accuracy in both clustering and rental income estimation using publicly available data, particularly for investment in rental properties, and can guide investors in making sound decisions in the real estate investment domain.

***Keywords*:** *Condominiums Rental Income (CRI), Data.gov Catalog Harvesting. New York State (DCH- NY), Z-score Transform (ZT), correlation coefficients (C-C), Fuzzy C-mean clustering (FCM), and linear regression (LR)*

Introduction

In today's world, the impact that specific real estate investments can have on the economy cannot be underestimated (Ambrose and Shen, 2023; Garriga and Tsouderou, 2023; Garriga et al., 2023; Lim, 2024; Aalbers et al., 2024); therefore, comprehending the complex nature of condominium rental income is imperative (Mehnaoui and Mehnaoui, 2019). Real estate investment is one of the most prominent methods for achieving long-term financial returns (Anderson et al., 2022). Developed markets exhibit passive renting and active buying (Liu et al., 2021). However, the return on real estate investments is crucial to financial decision-making and success (Robin, 2022). Furthermore, price returns are correlated with market rental growth (An et al., 2016). There are seven classical methods, including total return (Wang et al., 2022). This involves summing annual property revenue, including rents and proceeds, and dividing it by the total investment cost (Eichholtz et al., 2021; Saengchote & Charoenpanich, 2021). The positive relationship between performance of private equity real estate funds and voluntary reporting (Borysoff et al., 2024) of ESG disclosure and establishment that GRESB participation affects the real return (Gradillas et al., 2021), price change (Gopal & Pitts, 2025), and is unrelated to local economic factors (Devine et al., 2024). The study presents an innovative classification model and the capability to predict condominium rental income (Aleqab et al., 2025), thereby assisting investors with crucial insights into the market's dynamic facets to thrive (Behera et al., 2023). It has been reviewed that, as the condominium rental income becomes more important as a factor (Kinatta et al., 2021) of the real estate investment (Kinatta et al., 2021; Tamilmathi & Priya, 2023), there is still no satisfactory classification and prediction models of this income according to property market characteristics. Today's approaches tend to fail to capture interactions between the condominium market value (Potrawa & Tetereva, 2022), gross square footage (Tang et al., 2021) and total number of units which leads to undesirable investment decisions (Gattis, 2023) using machine learning in the evaluation of estate worth (Ho, et al., 2021; Steurer, et al., 2021). However, the study seeks to address the gap by developing a new model that employs Z-score transformation, correlation coefficients, Fuzzy C-mean clustering, and linear regression to provide a near-perfect measure of condominium rental income. Fernandez (2020) stated that company returns can be categorized into Total Return (TR) and Total Return for All Shareholders (TRAS) and observed that Companies in the S&P 100 index record annual TR exceeding 1% of their TRAS (Fernandez, 2020). Brounen et al. (2021) reported on 64 European public real estate companies from 2011 to 2018, with an average leverage of 40% and a sample return of 12.15% per year and focused on regional portfolios. Short-term real estate is more difficult to assess than larger, older estates, and both types are present in low- and high-income neighbourhoods (Bastos and Paquette 2024).

Net Cash Return (N-C-R) calculates the return on real estate investment (Agarwal et al., 2021; Walia et al., 2023). The net cash return calculation involves collecting all annual cash income from the property (McGaffin et al., 2019; Weber, 2021), such as rents for rental units and other revenues, and then deducting all expenses associated with it, such as maintenance (Airgood-Obrycki et al., 2023), management, and taxes (Robinson & Sensoy, 2016). Prior research by Aldboush et al. (2023) demonstrates that financial profitability determinants can be structured to improve predictive accuracy. The use of the income capitalization approach in the Enugu property market is therefore imprecise due to the market's forward-payment tenancy terms (Odimegwu & Anyakora, 2023). When we apply this approach to annual in-advance rental income cash flow, we obtain a higher capital value (Ifeanacho & Egbenta, 2023). Furthermore, ROI is a measure of real estate investment performance (Gigante & Cozzio, 2021), calculated by dividing the expected net return by the original investment cost and multiplying the result by 100 (Phillips, 2023; Riddiough, 2022). ROI measures investment effectiveness; a high ROI indicates greater attractiveness in real estate (Demirchelie, 2022). Realist research method uses quantitative (Mukumbang, 2023) and qualitative studies for deal evaluation (Farkas, 2023; Manzano, 2022). Moreover, it has been observed that estimating Future Value (EFV) is based on future revenue forecasts and accounts for the property's value appreciation (Figlioli & Lima, 2022) and urban expansion in the area surrounding the property. Future value estimation helps assess real estate returns and inform strategic investment decisions (Beracha & Downs, 2015; Steininger et al., 2021). Financial analysis uses the internal rate of return (IRR) and net present value (NPV) to assess the feasibility of real estate investments (Agung & Zuhri, 2023; Arjunan, 2022) and to determine whether to proceed with the investment (Tan, 2022). Ogunbiyi stated that efficiency cash-flow analysis models in Lagos, Nigeria, were developed to improve appraisal practice among Estate Surveyors and Valuers in line with property investment decisions (Ogunbiyi, 2023).

Moreover, Average Annual Return (AAR) is the average yield over a specified period (Jalihal Sharanappa, 2022), calculated as the total annual return divided by the number of years invested (Chen, 2024). Investment Opportunity Analysis (IOA) is a method used to evaluate the feasibility of real estate investments by analyzing expected returns, risks, income, expenses, inflation, market changes, and potential additional returns (Adam & Goyal, 2008)

The objectives of the study are to develop a robust framework for predicting CRI by leveraging key market features, including CMV, GGS, and TUN. Z-score transformation, correlation coefficients, fuzzy C-mean clustering, and linear regression techniques are employed. The applicability of the developed model is examined using data from the DCH-NY for 2008-2012, which may help potential real estate investors.

The study provides a clear model for accurately projecting rental income and enhancing decision-making for condominiums and rent control (Qiu, 2023). In addition to predictive accuracy, the study considers how varying levels of technological adoption influence platform profitability, market share, and service cost structure. The study employs statistical techniques, including Z-score transformation, fuzzy C-means clustering, and linear regression, to analyze factors in the condominium rental market. It helps policymakers understand the market and formulate sustainable growth policies. The paper introduces a new statistical model for predicting CRI data, thereby addressing a major problem in predicting real estate investment returns.

Literature Review

Over the last decade, studies have employed various statistical methods on datasets collected from diverse regions worldwide. These studies aimed to analyze and predict real estate investment returns by examining data from multiple contexts and markets (Kasim, et al., 2024). The analysis of previously published studies revealed significant challenges. Many studies relied heavily on specific factors, such as geographical topology (Mustofa et al., 2023) and transport services topology (Ifeanacho & Egbenta, 2023; Alnsour et al., 2024). Others focused on financial conditions (Mustofa et al., 2023) (Gichuru et al., 2022) or legal conditions of the studied area (Ngoc et al., 2023; Ahmed & Salam, 2022). These limitations restricted the generalizability of their findings. As a result, the findings were often limited to the context in which the data were collected and could not be generalized to other study cases. Some studies, such as those by (Marçal et al., 2023) and (Shen & Wilkoff, 2022), focused on the impact of epidemics and disasters on the CRI values of specific real estate properties (Shen & Wilkoff, 2022). However, these studies are limited to crisis periods and lack broader applicability. Additionally, most studies relied on classical statistical methods or theoretical strategies (Zhou, 2023). They typically began by determining descriptive statistics and then applied deterministic and restrictive solutions. Another issue is the substantial estimation error margins reported in previous research.

In many cases, the relationships between the studied features across datasets appeared nearly linear, which limits the accuracy of predictions. To address these challenges, this research proposes a novel statistical methodology to provide a more adaptive, simple, and efficient system for estimating condominium rental income (CRI) values (Alsaad & Al-Okaily, 2022). This approach uses independent and global features from related datasets to improve the realism and generalizability of the results (Al-Okaily, 2024). According to Almasria et al. (2024), organizational and technological innovation mediates the relationship between raw data inputs and sustainable performance outcomes.

The dynamics of the rental housing market have been the subject of extensive research across regions (Livingstone & Sanderson, 2022), aimed at understanding the factors influencing housing affordability, market trends, and the determinants of rental value. For instance, a study conducted by (Derkacz & Cohen, 2024) identified high flexibility in Poland's rental market trends, influenced by macroeconomic factors like the 2011–201g2 financial crisis (McGough & Berry, 2020) and the 2020 COVID-19 pandemic (Derkacz & Cohen, 2024). Using EU SILC surveys, their approach showed strong alignment with Polish statistical data. This validated the robustness of their evaluation methodology. Moreover, studies on rental housing dynamics highlight varied findings. For instance, (Benshak et al., 2024) linked rental value surges in Nigeria to tenant demographics, recommending housing supply policies (Benshak et al., 2024) (Adepetu, and Aliboh, 2024). (Menezes et al., 2023) emphasized the historical resilience of low-income rentals in Brazil (Menezes et al., 2023) (Lacerda and Primavera, 2023), while Budogo and Magina (2024) identified barriers to rental housing production in Tanzania, proposing solutions to address urban housing shortages (Budogo & Magina, 2024).

Building on these regional insights, Hełdak and Kucher (2024) found that location, residential standards, and property attributes influenced rental rates in Polish and Ukrainian border cities (Hełdak & Kucher, 2024). (Bello et al., 2020) identified construction costs, inflation, and neighbourhood characteristics as key determinants of residential property values in Abuja, Nigeria (Bello et al., 2020) (Adepoju, and Durosinmi, 2020). Similarly, (Madsen et al., 2022) demonstrated through machine learning that urban amenities significantly impact condominium prices in Colombo, Sri Lanka, highlighting the multifaceted drivers of real estate markets. There are different approaches to computing the return on a real estate investment (Pagourtzi et al., 2003; Himmelberg, et al., 2005; Fugazza et al., 2007; Hott & Monnin, 2008; Derkacz, 2024), including the following:

i. Total Return (TR): one of the most common ways to estimate the TR on real estate investment involves summing up all the annual revenue from the property, including rents for rental units and other proceeds such as the return on the sale of a property or increases in the value of the property over time, and then dividing it by the total investment cost.
ii. Net Cash Return (N-C-R): calculates the return on real estate investment. The net cash return calculation involves collecting all annual cash income from the property, such as rents for rental units and other revenues, and then deducting all expenses associated with it, such as maintenance, management, and taxes. The N-C-R is the actual income you can make from the property annually.
iii. Return on Investment (ROI) is an indicator for evaluating the performance of real estate investments. It is calculated by dividing the expected net return from the property by the original investment cost, then multiplying the result by 100 to obtain the percentage. ROI measures the

effectiveness of an investment and compares it to other investment opportunities. The high rate of return on investment is a strong indication of the attractiveness of real estate investment.

iv. Estimating Future Value (EFV): is based on future revenue forecasts and regards the increase in value of the property and urban expansion in the area surrounding the property. Future value estimation helps assess real estate returns and inform strategic investment decisions.

v. Financial Analysis (FA) includes the calculation of the internal rate of return (IRR) and net present value (NPV). The IRR is the rate of return on the present value of expected cash flows over the investment period. NPV is the value of all expected cash flows from an investment after being discounted at the appropriate discount rate. Financial analysis helps assess the feasibility of real estate investments and determine whether to proceed with the investment.

vi. Average Annual Return (AAR): estimates the average yield over a specific period, such as 5 or 10 years. The AAR is calculated as the percentage of the annual basis total return, and this ratio is then multiplied by the number of years to obtain the average return over the investment period.

vii. Investment Opportunity Analysis (IOA): is used to assess the feasibility of real estate investment by analyzing the expected returns and potential risks. This analysis includes estimating the property's expected income and expenses and assessing the associated investment risks. When using IOA, other factors should be considered, such as inflation, changes in the real estate market, and potential additional returns, including a long-term increase in property value.

These studies collectively highlight the complexity and variability of rental housing markets across different regions, emphasizing the need for context-specific policies and approaches to address challenges related to affordability, supply, and market dynamics (Madsen et al., 2022). Moreover, based on the foregoing literature review, two cross-cutting issues motivate our discussion of the comparative platform. These include: (i) data availability/standardization as a prerequisite for scalable analytics, and (ii) business-model design (monetization, barriers to entry) as a constraint on technology diffusion across markets. In the subsequent comparative analysis and case-based discussion, we revisit these themes by contrasting platform strategies across data-rich and data-constrained environments and linking them to observed scalability and monetization trade-offs.

Developed Model

The model was structured into four main parts. The first part comprises data preprocessing and the selection of numerical features that strongly correlate with CRI (AL-Khatib et al., 2024). Using cross-correlation among the numerical features in the original dataset, four features were selected that exhibited the highest correlation ratios: TUN, CGS, CMV, and CRI. After this, the model concluded the first part using the study's selected dataset, which contains 6083 records of the selected features.

The first part of the model provides an overview of the data. In contrast, the second part applies a statistical technique to estimate the number of categories or clusters into which the data may be divided. A developed technique is proposed that applies the Z-transform (ZT) to the features in the selected dataset, followed by representing the Z-distribution (ZD) and its correlation. This process yielded six z-scores with correlation coefficients near zero, indicating that the data could be partitioned into six clusters. In the third part of the model, the FCM technique is employed to cluster the selected dataset into the number of clusters identified in the second part. Six clusters are determined, along with their correlated centres. This developed model can efficiently classify any of the four entered features into one of the six clusters.

To perform CRI estimation, multiple experiments were conducted on the selected features to quantify the correlation between CRI and the other three features. Experiments included in this model resulted in a linear correlation between the calculated factor F2: (CRI per TUN) + (CRI per CGS), which is considered a dependent variable, and the calculated factor F1: (CMV per TUN) + (CMV per CGS), which is considered an independent variable.

Consequently, we observed that the linear correlation coefficients varied across the six features extracted in the model's third part. This motivated us to work in the fourth part of the model, using the LR technique to estimate the linear coefficients between F2 and F1 for each included cluster of data, and to finally estimate CRI accurately from three entered features: TUN, CGS, and CMV, after determining the appropriate cluster and fitting the LR model. Figure 1 illustrates the architecture of our developed model schematically.

To empirically validate the proposed model, a special-case study framework was employed. The companies were selected based on data disclosure, operational scale, and alignment with rent-based analytics. Specifically, Redfin was selected over Compass because it is both a real estate brokerage and an analytics platform that leverages technology to provide standardized, publicly available information on property values, unit features, and rental projections. Embedding transparent, auditable data streams—such as blockchain-verified transaction histories—reduces information asymmetry in predictive modeling (Almasria et al., 2026). This also makes Redfin particularly appropriate for modeling Condominium Rental Income (CRI) using quantitative methods. On the contrary, Compass does not provide the same level of publicly available, structured rental data needed for this study and primarily targets high-end residential markets and agent enablement. The geographical location was also selected to ensure complete data, a mature market with such data, and relevance to the existing literature, thereby supporting model validation and enabling extensive generalizability.

A) Part I

i. Data Preparation and Preprocessing:

The dataset used in this study was extracted from the Data.gov Catalog Harvesting (2008-2012) and contains condominium data for five New York cities during the four fiscal years from 2008 to 2012. The dataset includes a large number of records containing both numerical and non-numerical features. Due to duplicate observations for the studied cities across the selected fiscal years, a substantial number of duplicate records were identified. Additionally, some data entries were blank. During the preprocessing phase, these duplicate and incomplete records were removed. As a result, 6,083 unique records were retained for analysis.

ii. Feature Selection:

In this phase, cross-correlation was applied to the numeric features obtained from the previous preprocessing stage to identify the most relevant features that exhibited the strongest correlation with CRI, the target variable in this study. The final dataset comprises 6,083 records with four selected features: TUN,

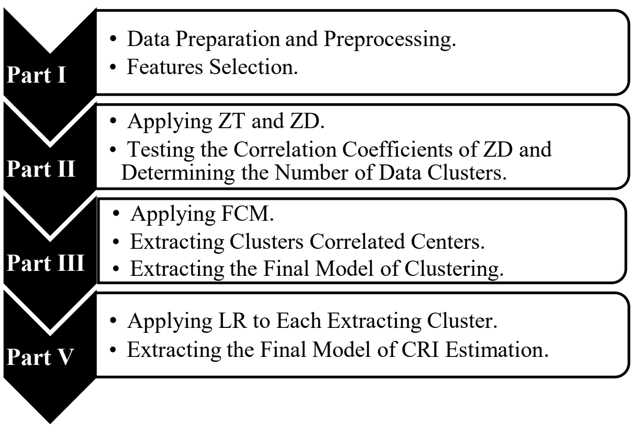
2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327

CGS, CMV, and CRI. These features show significant correlation coefficients, indicating robust linear relationships between each pair. Table 1 presents the correlation coefficients for the selected features of the constructed dataset. Due to the variability in feature value ranges, a normalization process was applied. For each feature value (FV), the normalized feature value (NFV) is shown in Equation 1.

$$CapNFV = \frac{\text{FV}}{\text{Max (FVs)}} \tag{1}$$

However, the normalization process ensures that all selected feature values fall within the range (Ahmed & Salam, 2022). This allows for a more straightforward representation of the dataset. Figure 2 illustrates a three-dimensional representation of the created dataset. The records in the dataset are sorted in ascending order based on the first feature (TUN).

Figure 1: The model architecture.



Note: A structured flowchart illustrating the five-step process of the model, beginning with data preparation and feature selection (Part I), followed by applying clustering techniques (Parts II & III), and finally regression modeling for CRI estimation (Part V). This diagram visually represents the logical flow of the methodology.

	TUN	**CGS**	**CMV**	**CRI**
TUN	1	0.963	0.612	0.708
CGS	0.963	1	0.737	0.821
CMV	0.612	0.737	1	0.971
CRI	0.708	0.821	0.971	1

PACIFIC RIM PROPERTY RESEARCH JOURNAL
2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327

Note: A Table 1: A correlation matrix showing the relationships among TUN, CGS, CMV, and CRI features. The high correlation values (e.g., CMV-CRI = 0.971) indicate strong dependencies between certain variables.

Availability and reliability of data are two determinants of differences among platform business models, as outlined in Table 1. Real estate data in Southeast Asian settings lack consistency, are fragmented, and lack verification; they are based on informal leasing systems, low disclosure requirements, and low levels of standardisation, which hinder the implementation of analytics-powered channels. As a result, companies would prefer to use agent-based listing/advertising/lead-generation models rather than data-intensive predictive systems. The business models have intrinsic limitations: agent-based and listing platforms have weak scalability and modest predictive power due to the limited depth of their data; analytics-based platform types - common in data-rich markets - require considerable fixed investments in data infrastructure, regulatory compliance, and model development; iBuyer platform has steep capital requirements and is dependent on volatile price patterns; and blockchain platforms have high uncertainty costs, high implementation costs, and poor adoption in informal and low-value markets.

Moreover, the correlation matrix in Table 1 reveals distinct performance-related patterns across different levels of technological adoption. More importantly, there are strong correlations between CMV and CRI ($r = 0.971$) and between CGS and CRI ($r = 0.821$). These data points indicate a scalable increase in rental income with respect to asset size and market value in data-rich environments. The results indicate an increase in revenue-to-asset density, a key measure of relative profitability, among platforms that utilize advanced analytics, automated valuation, and data-driven pricing systems. The smaller correlations with the TUN, by contrast, refer to ecosystems in which income production is less efficiently captured and is generally associated with lower technological integration. The fact that robust dependencies are more prevalent than other high-value features indicates that market structures characterized by sophisticated analytical implementation can derive greater value from each transaction and achieve a superior market position. Consequently, Table 1 provides empirical evidence that increased technological use is associated

2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327

with higher income and market-share concentration, whereby positioning platforms that have mastered the structural associations take control of the high-end rental market.

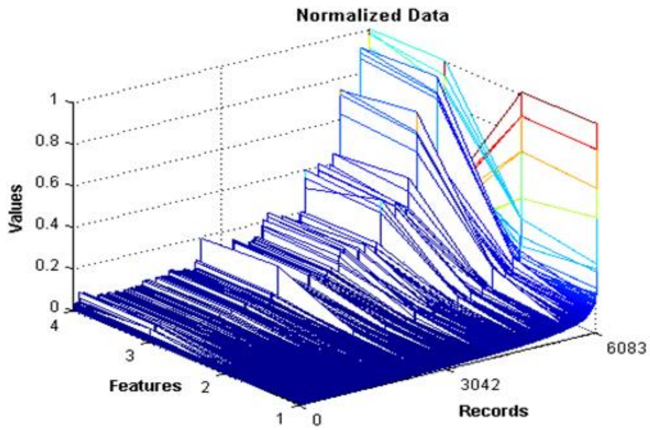


Figure 2: 3D representation of the dataset.

Note: A 3D graph visualizing the dataset with normalized values on the Y-axis, features on the X-axis, and record indices on the Z-axis. The colored lines depict variations across dataset records, illustrating the distribution and density of the data points.

B) Part II

i. Applying ZT and ZD:

A Z-score describes the position of a value in raw data (a feature in this case) in terms of its distance from the mean, measured in standard deviation units. It allows comparisons of scores across samples, even when their means and standard deviations differ. The Z-score was calculated using Equation 2. Figure 3 illustrates the calculated Z-scores for each studied feature.

$$Z_i = \frac{FV_i - Mean(FV)}{STD(FV)} \tag{2}$$

Using the calculated Z-score vectors, the corresponding ZD for each vector was computed. ZD represents the relative histogram of Z-score values. Figure 4 shows the resulting ZD of each studied feature.

PACIFIC RIM PROPERTY RESEARCH JOURNAL
2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327

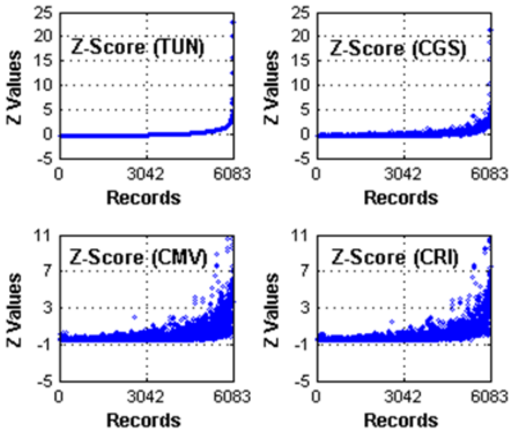


Figure 2: Z-score of each studied feature

Note: A set of four scatter plots showing Z-score distributions for different features (TUN, CGS, CMV, CRI). Each plot represents how feature values deviate from the mean, identifying outliers and data spread.

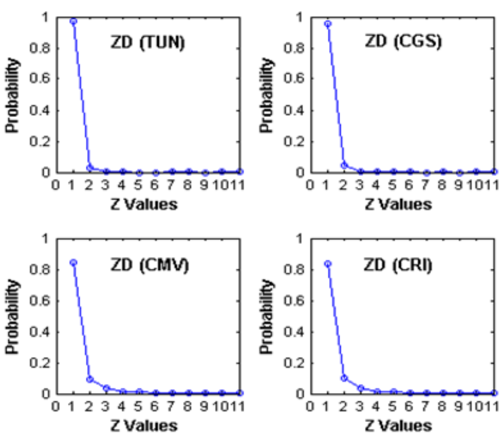


Figure 3: ZD of each studied feature

Note: Four probability distribution graphs of ZD values for each feature (TUN, CGS, CMV, CRI). The rapid decline in probability suggests a concentration of values near lower ZD ranges, indicating potential data skewness.

In this stage, the correlation coefficients (r) between the ZD vectors extracted in the previous stage were calculated using Equation 3. The correlation coefficients indicate the strength of the linear relationship between the ZD vectors; low r values indicate a weak linear relationship. Experiments showed that $r < 0.25$ indicates a weak linear relationship between the ZD vectors. Based on this, the number of Z-values in ZD

PACIFIC RIM PROPERTY RESEARCH JOURNAL
2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327

meeting the condition $r \leq 0.54$ was extracted. This number represents the number of possible clusters into which the dataset can be partitioned. The number of clusters is determined to be six. Figure 5 illustrates the results of the correlation coefficient calculations for the ZD vectors.

$$r = \frac{n(\sum XY)-(\sum X)(\sum Y)}{\sqrt{[n\,(\sum X^2-(\sum X)^2][n(\sum Y^2-(\sum Y)^2)^2]}} \tag{3}$$

Where r is the correlation coefficient and n: is the number of values in both X and Y vectors.

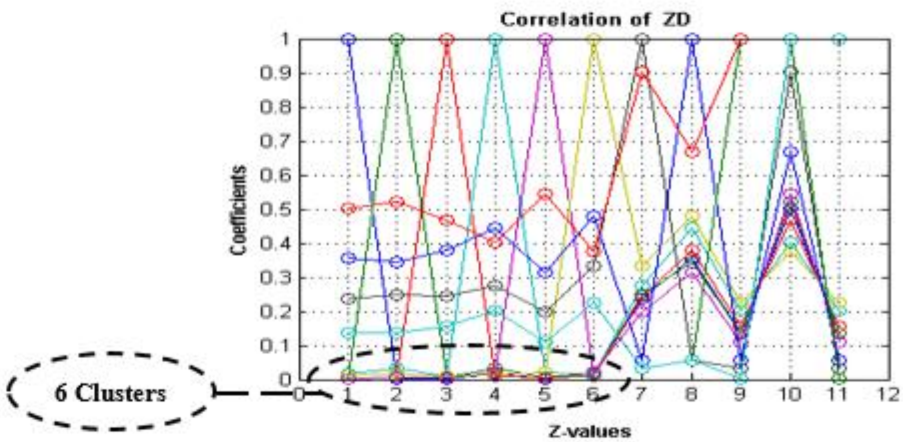


Figure 4: Correlation coefficients of ZD vectors

Note: A multi-line graph representing the correlation coefficients of different ZD vectors across Z-values. The six detected clusters are highlighted, showing how ZD values form distinct groupings within the dataset.

i. Applying FCM

Based on the specified number of clusters from the previous stage, FCM (Szilágyi et al., 2020) first determines initial centroids for each estimated cluster, with each cluster containing four vectors, corresponding to the number of selected features. FCM then calculates the distances between all centroids and each record in the dataset. Each record is assigned to one of the six clusters if it realizes the minimum distance to its centroid, as described in Equation 4. FCM updates the centroids by computing the mean of the records in each cluster. During this process, FCM computes the clustering error, defined as the sum of squared distances between the cluster centroids and the records in each cluster, as shown in Equation 4. FCM iterates in this manner until either a prespecified error threshold is met or a specified number of iterations is reached. The FCM was executed for 100 iterations, and the minimum error achieved ensured accurate performance. Figure 6, Sections A and B, illustrates the extracted clusters and their correlated centres. Figure 7, Sections A and B, present the distribution of clusters in the study dataset and the cluster values, respectively. Table 2 provides further clarification on both the cluster distribution and descriptions.

The SWOT matrix of real estate platforms highlights key internal and external dynamics, revealing that barriers to market entry and monetization models significantly affect strategic insights. Geographic challenges, such as regulatory compliance and data accessibility, affect Southeast Asian platforms, which struggle with fragmented data and informal transactions, whereas U.S. platforms leverage MLS systems for scalable monetization. Additionally, alternative technologies, such as iBuyers and blockchain, face barriers, including high infrastructure costs and legal hurdles. Monetization models vary, with examples including Redfin's integrated services and Southeast Asian platforms that focus on agent-paid listings and advertising (Brij, 2025). Understanding these models is crucial for identifying strategic opportunities and threats in diverse markets.

$$over\ i\mathrm{D_k} = \textstyle\sum_i (\mathrm{R_i} - \mathrm{C_k})^2 \tag{5}$$

$$\mathrm{E_{rr}} = \textstyle\sum_k \sum_i \mathrm{D_{ki}} \tag{6}$$

Table 2: Clusters description and distribution (L, M, and H refer to low, medium, and high values of the studied features, respectively.

		Features				**Distribution %**
		TUN	**CGS**	**CMV**	**CRI**	
Clusters	**1**	M	M	H	H	2.17
	2	L	L	L	L	62.06
	3	L	M	M	M	22.49
	4	M	M	M	M	9.19
	5	M	M	H	M	3.78
	6	H	H	H	H	0.31

Note: Table 2: A tabular classification of six clusters based on low (L), medium (M), and high (H) values for each feature. The majority of the dataset belongs to Cluster 2 (62.06%), representing low feature values, whereas Cluster 6 (0.31%) contains high-value records.

In addition to presenting the statistical distribution of condominium clusters, Table 2 highlights a set of frameworks for conducting a rigorous assessment of strategic entry and monetization obstacles in the current context of real estate analytics applications. The dominance of Cluster 2 (62.06) - typified by low TUN, CGS, CMV, and CRI - stands out as a defining feature of markets in which rental assets remain small-scale, fragmented, and yield lower returns than those in larger markets. Under these conditions, typical of emerging economies across Southeast Asia, informal leasing agreements, a lack of data standardisation, and insufficient regulatory transparency hinder entry to platforms. This forces platforms to implement

PACIFIC RIM PROPERTY RESEARCH JOURNAL
2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327

low-margin monetisation models, including agent-paid listings and advertising-driven models, thereby limiting the application of advanced analytics.

On the other hand, medium- to high-value clusters (Clusters 1, 5, and 6), albeit with a smaller proportion, represent the high-value market segments, in which entry barriers are more technological than structural. Such clusters are associated with mature economies, most prominently the United States where the standardised data infrastructure (e.g., MLS systems), the formalised valuation models and the definite regulatory frameworks play a role in keeping the costs of data acquisition low and allow the economies of scale to be acquired in the platform based on analytics. Examples of this model include firms such as Redfin, which monetize built-in business models based on a combination of brokerage commissions, premium analytics services, and investor-oriented decision-support tools, thereby directly transforming data insights into real economic value.

The emergence of alternative technology platforms evidences the introduction of additional layers of entry barriers. iBuyer-type models are characterized by substantial capital requirements and price volatility, which makes them less viable in clusters with stable, high-CRI values. The promises of increased transparency and transactional efficiency do not dispel legal uncertainties or the lack of interoperability, making blockchain-based real estate platforms inaccessible to many users, especially in clusters where property values are low or informal. This means that such innovations are likely to remain confined to niche or experimental markets.

It is important to note that the clustering framework identified in Table 2 underscores that monetisation strategies must be consistent with the typology of property values within each cluster. Clusters of high value warrant premium, data-intensive, revenue models targeting institutional investors and developers, low-value clusters require volume-based or intermediary-mediated monetisation. Table 2 is a bridge-analytical element; it converts the technical results of the clustering into a practical model of evaluating opportunities and threats in different geographic areas and technological systems. This would provide both scholars and practitioners with an overall view of how to align analytical sophistication with market

2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327

realities.

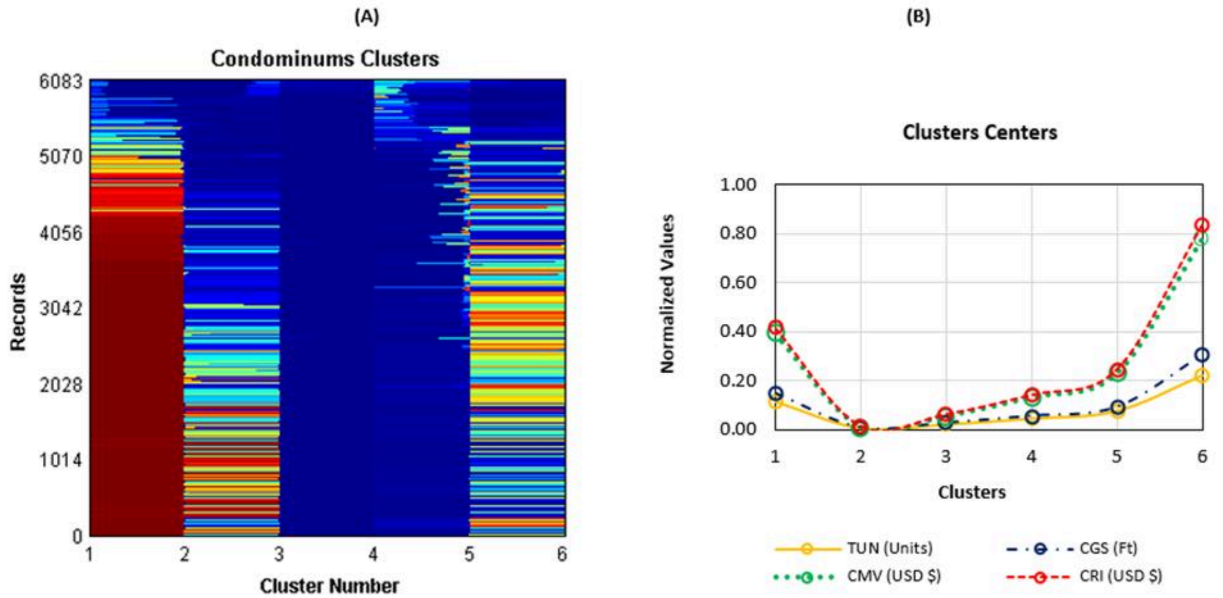


Figure 5: A: The extracted clusters and B: The clusters correlated centres

Note: (A) A heatmap showing the classification of records into six clusters based on feature similarities.

(B) A line graph representing the cluster centers, displaying the normalized values of features (TUN, CGS, CMV, CRI) across six clusters.contains high-value records.

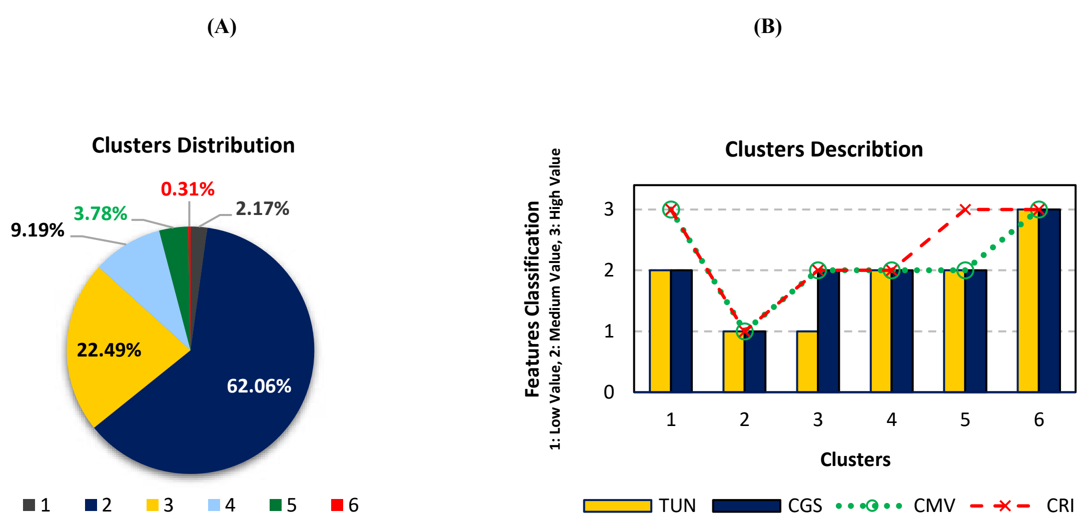


Figure 6: A: Clusters distribution and B: Clusters description among the studied features.

Note: Cluster 6 is the smallest (0.31%).
(B) A bar chart classifying the clusters into low (L), medium (M), and high (H) feature values, helping in cluster interpretation across six clusters that contains high-value records.

PACIFIC RIM PROPERTY RESEARCH JOURNAL
2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327

Part IV

i. Applying LR

In this stage, a linear correlation between two factors was calculated. The first factor, F1, is calculated from CMV, TUN, and CGS, and the second factor, F2, is calculated from CRI, CMV, and CGS, as shown in Equations 7 and 8, respectively. This correlation varies in coefficient values across clusters.

$$\mathrm{F1} = \frac{\mathrm{CMV}}{\mathrm{TUN}} + \frac{\mathrm{CMV}}{\mathrm{CGS}} \quad (7)$$

$$\mathrm{F2} = \frac{\mathrm{CRI}}{\mathrm{TUN}} + \frac{\mathrm{CRI}}{\mathrm{CGS}} \quad (8)$$

LR was applied to each cluster separately to determine the adaptive linear model for each dataset cluster (Marçal et al., 2023). LR begins to give an initial weight (w_0, and w_1) to the hypothesis illustrated in Equation 9 below, and uses the mean squared error function MSE to calculate the error between the real F2 values and the estimates by the linear hypothesis, see Equation 10

$$\mathrm{h}\ (\mathrm{F1})_{\mathrm{w}} = \mathrm{w}_0 + \mathrm{w}_1\ \mathrm{F}_1 \quad (9)$$

$$\mathrm{MSE} = \frac{1}{2\,\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} (\mathrm{h}^{\mathrm{i}} - \mathrm{F2}^{\mathrm{i}})^2 \quad (10)$$

Where n is the number of records in each cluster.

Then LR iterates to minimize MSE and update the weights w0 and w1 by applying the gradient descent of MSE on each weight according to Equations 11 and 12, respectively:

$$\mathrm{w}_0^{\mathrm{k}+1} = \mathrm{w}_0^{\mathrm{k}} - \frac{\propto}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} (\mathrm{h}^{\mathrm{i}} - \mathrm{F}_2^{\mathrm{i}})$$

(11)

$$\mathrm{w}_1^{\mathrm{k}+1} = \mathrm{w}_1^{\mathrm{k}} - \frac{\propto}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} (\mathrm{h}^{\mathrm{i}} - \mathrm{F}_2^{\mathrm{i}}) \mathrm{F}_1^{\mathrm{i}}$$

(12)

Where α: is a constant that refers to the learning rate, it is chosen to be a small value to control the step in the MSE between two sequential iterations to be stored to non-jump in the max range of values over the local minima.

LR is trained in a reasonable number of iterations varying from 20 to 35 over each cluster of data and stabilizes its coefficients (w_0, and w_1) to represent a robust linear correlation between the studied factors (F2 as a dependent and F1 as an independent variable). Table 3 reports the results of applying the LR model over the variables of each extracted condominium cluster. The LR coefficients represent the results of this stage w_0, and w_1 According to Equation 9, the R-squared value shows how well the data fit the LR model, the ideal R is close to 1, and the root MSE measures how far from the regression line data points are, standard error represents the average distance between the observed values and the regression line, and P > absolute value tests the null hypothesis that the independent variable does not correlate with the dependent variable. If the p-value is less than a specified significance level (e.g., 0.05), the predictor variable is considered statistically significant in the model. Figure 8 illustrates the resulting linear model

2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327

over the specified clusters in this study. The trained LR model was tested on the studied data by entering three features, TUN, CGS, and CMV, to calculate the F1, and the LR model estimates F2, then it estimates CRI by using the entered TUN and CGS, the trained LR demonstrated a high accuracy in estimating correctly the real CRI, with the significant part of error values are close to zero. Figure 9 illustrates the estimation performance of LR. Figure 10 illustrates the estimation error histogram.

Table 3: The results of applying LR on the extracted variables F1 and F2 from the studied condominium clusters.

Clusters	Observations	Coefficients		R-squared	Root MSE	Std. Err	P>|t|
		w_0	w_1				
I	559	0.720	0.935	0.956	0.516	0.008	0.000
II	1368	0.922	0.911	0.935	0.785	0.006	0.000
III	19	0.695	0.971	0.991	0.608	0.022	0.000
IV	3775	0.597	0.989	0.969	0.423	0.003	0.000
V	230	0.798	0.921	0.967	0.482	0.011	0.000
VI	132	0.685	0.950	0.958	0.645	0.017	0.000

Note: Table 3: A statistical summary of linear regression results for different clusters, including coefficients, R^2 values, root mean square error (RMSE), and significance levels (p-values). The high R^2 values (above 0.95) confirm the strong predictive power of the model.

PACIFIC RIM PROPERTY RESEARCH JOURNAL
2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327

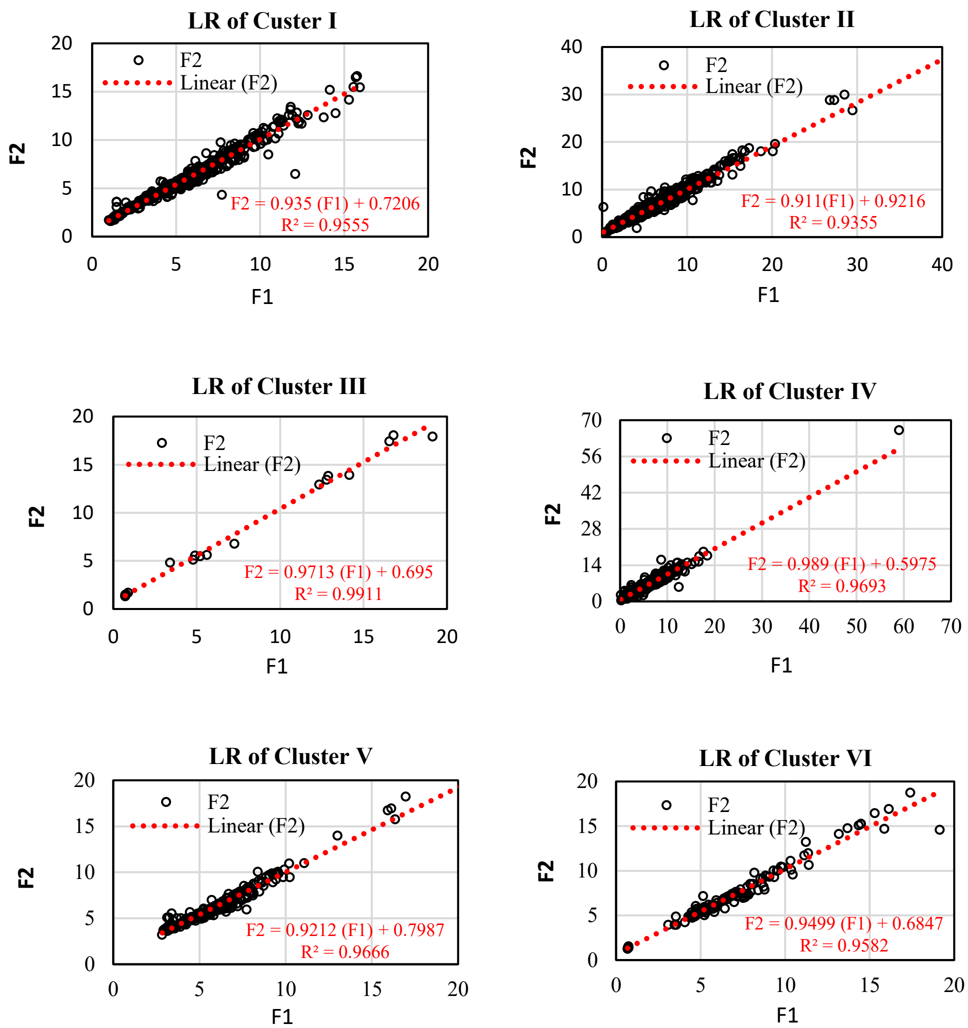
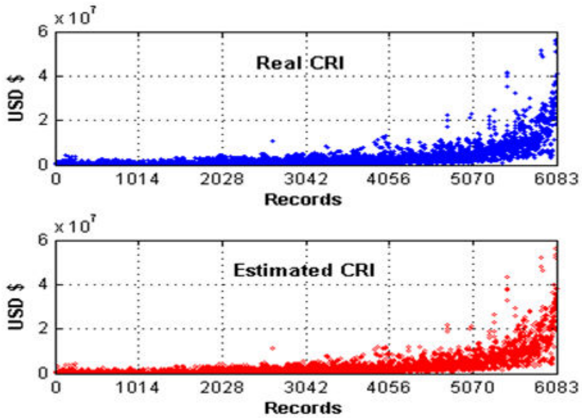


Figure 7: LR model over the different studied clusters of data.

Note: Six scatter plots showing linear regression (LR) models applied to different clusters. Each plot represents a specific cluster's F1 (independent variable) and F2 (dependent variable) relationship, with regression lines and R^2 values, indicating the model's predictive strength.

2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327



Note: Two scatter plots comparing actual CRI values (blue) and estimated CRI values (red). The similarity in trends between real and predicted values validates the LR model's accuracy.

Figure 8: The performance of LR in estimating CRI.

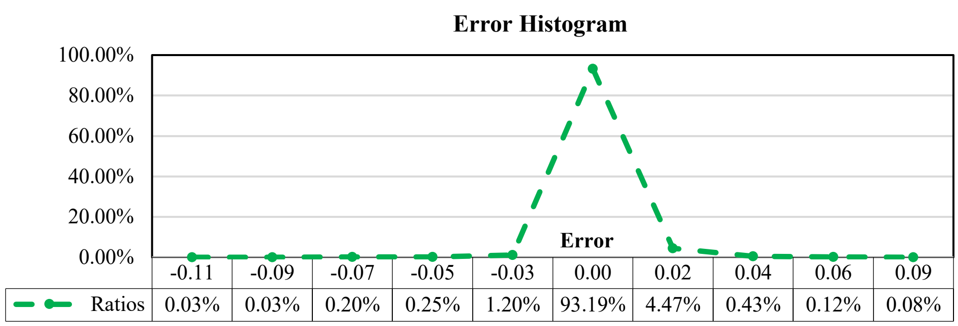


Figure 9: The estimation error histogram.

Note: A histogram displaying the error distribution in CRI estimation. The majority of errors cluster around zero (93.19%), indicating a highly accurate estimation model.

Implications

The main challenge in ensuring the success of the developed system, as well as similar systems discussed in the scientific literature (Benshak et al., 2024; Derkacz & Cohen, 2024; Madsen et al., 2022; Mustofa et al., 2023), is the linear-like relationship between the independent numerical variables in the relevant databases and the dependent variable, CRI. This relationship often leads to errors in estimation when applying a standard linear model. To address this challenge, the proposed framework first classifies properties into homogeneous clusters, allowing more appropriate linear relationships to be modeled within each class.

2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327

From a managerial perspective, this clustering structure provides a practical basis for decision-making related to market entry, pricing, monetization, and technology investment. By identifying clusters with similar rental-income behavior, platform managers and developers can align analytical complexity and operational strategies with the underlying structure of the market rather than applying uniform solutions across heterogeneous environments. These are strategic considerations based directly on the cluster structure and explain why the proposed model not only helps estimate the rental income but also informs platform-level decisions related to entering, monetizing, and positioning at competitive positions in different geographic and technological settings. One such example is the limitations of scalability caused by a geographically limited focus, i.e., the iProperty.com.my platform. The value of the platform, in that case, is directly connected to the presence of the local data, existing regulatory frameworks, and manual verification systems relying on agent-created listings and advertisement-based revenue models. These findings indicate that geographic expansion, without simultaneous investment in data standardization and regulatory alignment, is bound to limit cross-regional network effects and hinder long-term growth.

Additionally, it should be mentioned that the comparative price and cost of services charged by the other real-estate platform business models are virtually ignored in most studies, including previous implementations of predictive models. The fixed costs of analytics-based platforms are higher as they invest in data collection, computing infrastructure, model development, and specialized human capital, which they recoup via commission-based pricing, subscriptions, or offering bundled services. Compared to agent-based and listing-based platforms, agent-based platforms are lower-tech and operationally lean and follow flat listing or pay-per-lead-based models, both of which reduce entry barriers but constrain intensity and scalability of revenues. iBuyer models are capital-intensive and have high liquidity and risk-management costs, whereas blockchain-based platforms are expensive to implement and comply with. These cost considerations and predictive performance can provide a deeper insight into platform competitiveness, profitability, and strategic positioning.

Traditionally, high returns have been the primary motivation for real estate investments (Ngoc et al., 2023). Providing investors with a clear understanding of property categories based on size (area, total units), market value, and potential gross rental income can increase confidence in investment decisions. Integrating Big Data from transaction records and macroeconomic indicators significantly improves rental yield estimates (Almasria et al., 2024). The developed model provides investors, governmental or private, in regions similar to the one used in this case study, with an estimated rental income for a specific property based on simple and globally applicable features.

This study contributes significantly by demonstrating the model's capacity to classify real estate properties effectively and estimate rental income with a high degree of accuracy. The model's simplicity and global applicability make it a valuable tool for investors in various regions. Additionally, it offers a practical approach that can be adapted to worldwide real estate markets, providing essential insights for seasoned and new investors.

While the findings are promising, further research could explore how the model adapts to varying market conditions and the potential for refining classification methods to account for other variables not considered in this study. Future studies could also evaluate the model's performance in different regions and with different real estate types, further validating its robustness and generalizability.

PACIFIC RIM PROPERTY RESEARCH JOURNAL
2026, VOL.31, No.1, 53 – 78
DOI: https://doi.org/10.64361/14445921.2026.212327

Discussion and future work

The constructed model shows good predictive research in approximating the condominium rental income (CRI) through the important market characteristics- CMV, CGS, and TUN, and is reliable in clustering properties into fuzzy C-means and employing specific linear regression as per each category (Alkhwaldi et al., 2024). But, in addition to its technical correctness, the model has important implications for managers, which should be noted.

The model is a useful tool for investors to help them estimate the expected income prior to making investment decisions, thus being able to risk profile and diversify their portfolios better. This is consistent with Pisal et al. (2025), who stressed goal-based portfolio sifting on the basis of risk tolerance. The same report done by Torkian et al. (2025) established that AI-based models are useful in balancing high-risk investments with diversified portfolios. Prioritizing CRI, the model aids in pre-investment forecasting and the reduction of exposure to poorly performing assets, which is important in multi-objective decision-making (Zheng et al., 2025). It can improve the real estate investment strategies, which have historically lacked by uncertainty and the inability to integrate data (Ntabanganyimana et al., 2025; Ibrahim et al., 2025). This indicates that the model enhances the accuracy of analysis and the belief of investors in real estate markets. The developers and property managers can use the knowledge of clustering to match property specifications to ranges of properties, which helps develop property planning and pricing strategies. Clustering used in the model helps developers and property managers to obtain practical information to match the property specification and the market segments on an income basis in order to improve the planning and pricing strategies (Al-Bashayreh et al., 2022). The clustering techniques enable stakeholders to derive data patterns and customize the development of urban areas more specifically, especially when the urban environment is mixed (Al-Rimawi & Nadler, 2025; Almasria et al., 2025; Алмасрія et al., 2025). This can be used by developers to maximize the floor plans, amenities, and unit sizes based on the estimated ranges of rental incomes to reduce the risk of over- or underpricing (Wong et al., 2025). Furthermore, clustering has been found to be helpful in property classes and valuation behaviour differentiation that could be used in the development of differentiated pricing in various neighborhoods (Kasim et al., 2025; Ibrahim et al., 2025). These strategies help in more strategic planning of the portfolios, and less risk in market entry (Alsmadi et al., 2023). Thus, the combination of clustering and real estate analytics fuses the field of data science with spatial development decision-making- creating value-aligned planning intelligence.

The capacity of the model to find clusters of neighborhoods is a useful resource to the policy-makers who want to use it to take specific action, like providing affordable housing or zoning policy changes. Decision-makers can find it easier to distribute resources and plan urban development strategies by mapping areas with similar features, such as low-income density or high pressure on renting in them. Clustering has also successfully been used to identify vulnerable slum regions and city risks, among other related applications, to develop housing and infrastructure planning (Bante et al., 2025; do Nascimento et al., 2025) (do Nascimento et al., 2025; Bante et al., 2025). Socio-spatial inequalities are also measured with the help of models and contribute to more equal urban development (Chen et al., 2025). Moreover, the incorporation of clustering into predictive planning does not go against the increasing popularity of AI-based urban policy instruments (Zhang & Zhang, 2025). Therefore, the model will bridge data analysis and governance, which will allow making policymaking more precise and meaningful.

To improve predictive modeling in real estate, it should be integrated into user-friendly, low-code platforms, supported by workshops from real estate associations. This integration addresses generational and technological shifts, offering actionable recommendations based on data analysis. A case study confirms the model's accuracy, with estimated rental income closely matching actual data. Future research should involve testing with diverse datasets to enhance generalizability and adaptability, potentially by adding dynamic features and exploring new applications in real estate analytics.

Conclusions

The integration of machine learning with time series analysis enhances cryptocurrency forecasting by minimizing risk and increasing trust. This study combines econometric and advanced machine learning models, overcoming the shortcomings of deterministic models in real estate analysis. It introduces an improved statistical model that employs global data characteristics and techniques like Z-transform, Z-distribution, and Fuzzy C-Means clustering to estimate condominium rental income, delivering accurate insights for global real estate markets. Multi-factor determinants, innovation-driven intermediation, and auditable data integration produce dynamic, evidence-based income predictions (Aldboush et al., 2024; Alduais et al., 2022).

This model significantly improves predictive risk assessment, providing essential tools for investors, policymakers, and analysts in the dynamic global real estate landscape. The practical findings of NYC data corroborate the legitimacy of the issues expressed in the previous studies and demonstrate a distinct course of action where machine learning applications are applied not only to predict, but also to gain a strategic understanding of the real estate investment and policy. Implementing this model in real estate decision-making, such as urban planning and investment management, can lead to advanced, data-driven analytics. Its value in policy and investment shows that practical innovation and theoretical critique can and should complement each other.

References

Aalbers, M. B., Taylor, Z. J., Klinge, T. J., & Fernandez, R. (2023). In real estate investment we trust: State de-risking and the ownership of listed US and German residential real estate investment trusts. Economic Geography, 99(3), 312-335.

Adam, T., & Goyal, V. K. (2008). The investment opportunity set and its proxy variables. Journal of Financial Research, 31(1), 41-63.

Agarwal, D., Bansal, S., Jain, A., & Jain, S. (2021). A Study on Investment Behavior of Individual Investors With Reference To Delhi-NCR. International Journal of All Research Education and Scientific Methods (IJARESM), 9(9).

Agung, T. S., & Zuhri, B. S. S. (2023). Analysis of the Financial Feasibility of Potential Post-Pandemic Businesses Using the Net Present Value (NPV), Internal Rate of Return (IRR), and Payback Period (PP) Methods (Case Study: MSME Environmentally Friendly Bioplastic Products). Jurnal Multidisiplin Madani, 3(7), 1432-1441.

Ahmed, S., & Salam, M. (2022). Rental housing policies and associated legal covers: Case of middle-income formal housing in Karachi. Journal of Urban Management, 11(4), 488-499.

Airgood-Obrycki, W., Hermann, A., & Wedeen, S. (2023). “The rent eats first”: rental housing unaffordability in the United States. Housing Policy Debate, 33(6), 1272-1292.

Al-Rimawi, T., & Nadler, M. (2025). Leveraging smart city technologies for enhanced real estate development: An integrative review. Smart cities, 8(1), 10.

Aleqab, M., Airout, R., Alawaqleh, Q., & Almasria, N. (2025). The mediating role of innovation in the relationship between corporate governance and organisation performance. International Journal of Business Information Systems, 49(3), 431-444.

Almasria, N. A., Bataineh, S., Aldboush, H., & Almajali, A. (2025). Investigating the role of CEO characteristics and environmental factors on cloud accounting information system. Journal of Information Systems Engineering and Management. 2025,10(47s). Available: https://jisem-journal.com/index.php/journal/article/view/9221/4268.

Алмасрія, Н. А., Ершаїд, Д., Джалгум, Я. А., & Алмаджалі, А. (2025). The role of fintech in transforming risk management and financial services: A systematic review and meta-analysis. Financial and credit activity problems of theory and practice, 2(61), 409-429.

Anderson, R., Beracha, E., & Propper, S. (2022). The impact of real estate allocation on investors’ ability to generate real income. Journal of Property Research, 39(2), 120-147.

Aldboush, H. H., Almasria, N. A., & Ferdous, M. (2023). Determinants of firm profitability: empirical evidence from Jordan’s service sector. Business: Theory and Practice, 24(2), 438-446.

Aldboush, H., Almasria, N. A., Tawfiq, T. T., & Tawaha, H. (2024). The nexus between dividend payout ratios, return on assets and leverage: A 15-year analysis of fortune 500 companies. International Journal of Economics and Finance Studies, 16(4), 40–73.

Almasria, N. A., Alhatabat, Z., Ershaid, D., & Baarimah, A. O. (2026). Integrating Blockchain and Big Data into External Audit Process: Perspective and Challenging. In Generative Artificial Intelligence-Driven Innovation: Strategies and Ethical Frameworks for Responsible AI (pp. 491-509). Cham: Springer Nature Switzerland.

Almasria, N. A., Alhatabat, Z., Ershaid, D., Ibrahim, A., & Ahmed, S. (2024). The mediating impact of organizational innovation on the relationship between fintech innovations and sustainability performance. Sustainability, 16(22), 10044.

Алмасрія, Н. А., Ершаїд, Д., Джалгум, Я. А., & Алмаджалі, А. (2025). The role of fintech in transforming risk management and financial services: A systematic review and meta-analysis. Financial and credit activity problems of theory and practice, 2(61), 409-429.

Arjunan, K. (2022). A new method to estimate NPV and IRR from the capital amortization schedule and the advantages of the new method. Australasian Accounting, Business and Finance Journal, 16(6), 23-44.

Al-Okaily, M. (2024), Does AIS usage matter in SMEs performance? an empirical investigation under digital transformation revolution. Information Discovery and Delivery, 52 (2): 125–137. https://doi.org/10.1108/IDD-08-2022-0072

Alsaad, A., & Al-Okaily, M. (2022). Acceptance of protection technology in a time of fear: the case of Covid-19 exposure detection apps. Information Technology & People, 35(3), 1116-1135.

AL-Khatib, A. W., Shuhaiber, A., Mashal, I., & Al-Okaily, M. (2024). Antecedents of Industry 4.0 capabilities and technological innovation: a dynamic capabilities perspective. European Business Review, 36(4), 566-587.

Alkhwaldi, A. F., Al-Qudah, A. A., Al-Hattami, H. M., Al-Okaily, M., Al-Adwan, A. S., & Abu-Salih, B. (2024). Uncertainty avoidance and acceptance of the digital payment systems: a partial least squares-structural equation modeling (PLS-SEM) approach. Global Knowledge, Memory and Communication, 73(8-9), 1119-1139.

Al-Bashayreh, M., Almajali, D., Al-Okaily, M., Masa'deh, R. E., & Samed Al-Adwan, A. (2022). Evaluating electronic customer relationship management system success: the mediating role of customer satisfaction. Sustainability, 14(19), 12310.

Alsmadi, A. A., Al-Okaily, M., Alrawashdeh, N., Al-Gasaymeh, A., Moh'd Al-hazimeh, A., & Zakari, A. (2023). A bibliometric analysis of green bonds and sustainable green energy: evidence from the last fifteen years (2007–2022). Sustainability, 15(7), 5778.

Alnsour, J., Arabeyyat, A. R., Alnsour, A. J., & Almasria, N. A. (2024). The Impact of Financial Development, Foreign Direct Investment, and Trade Openness on Carbon Dioxide Emissions in Jordan: An ARDL and VECM Analysis Approach. Journal of Risk and Financial Management, 17(11), 490.

Bante, A. B., Dararo, K. F., Taddesse, N. K., & Dejene, I. N. (2025). Geospatial mapping and prediction of informal settlement growth in Addis Ababa using GIS and logistic regression model. Discover Environment, 3(1), 259.

Behera, I., Nanda, P., Mitra, S., Rath, S., & Senapati, S. (2023). Analyzing Market Volatility: Big Data's Insight into Macroeconomic Indicators. 2023 International Conference on Data Science & Informatics (ICDSI),

Bello, N. A., Adepoju, A. S., & Durosinmi, W. A. (2020). Factors influencing rental and capital values of residential investment property in Abuja, Nigeria. International Journal of Real Estate Studies, 14(1), 102-109.

Benshak, A. B., Adepetu, A. N., & Aliboh, V. (2024). Assessment of the Variations in Residential Rental Values in Rantya-Gyel Jos, Plateau State, Nigeria. African Journal of Environmental Sciences and Renewable Energy, 14(1), 13-29.

Beracha, E., & Downs, D. H. (2015). Value and momentum in commercial real estate: A market-level analysis. Journal of Portfolio Management, 41(6), 48.

Borysoff, M. N., Mason, P., & Utke, S. (2024). Understanding private equity funds: A guide to private equity research in accounting. Journal of Financial Reporting, 9(1), 21-49.

Brij, V. K. T. (2025). The Monetization Playbook: Digital Transformation Success in the Digital Economy. Journal of Computer Science and Technology Studies, 7(4), 116-124.

Budogo, B. N., & Magina, F. B. (2024). Access to Land, Building Materials & Finance for Rental Housing Production in Dodoma Capital City, Tanzania. International Journal of Social Science Research and Review, 7(3), 52-65.

Chen, B. J. (2024). Average Annual Return (AAR): Definition, Calculation, and Example. Investopedia. https://www.investopedia.com/terms/a/aar.asp

Chen, M., Robinson, C., & Singleton, A. (2025). Mapping multidimensional energy deprivation: Socio-spatial inequalities and policy implications in Great Britain. Computers, Environment and Urban Systems, 121, 102324.

Data.gov. (2025). Catalog Harvesting, New York State (NY) (2008-2012). Available at: https://catalog.data.gov (Accessed: 13 January 2025).

Demirchelie, E. (2022). Visual Communication of Key Concepts in Commercial Real Estate Analysis and Investment Massachusetts Institute of Technology].

Derkacz, A. J., & Cohen, V. (2024). The Size of the Rental Housing Segment in Poland and its Main Determinants. Real Estate Management and Valuation, 32(1), 49-57.

Devine, A., Sanderford, A., & Wang, C. (2024). Sustainability and private equity real estate returns. The Journal of Real Estate Finance and Economics, 68(2), 161-187.

do Nascimento, G. A., Giannotti, M., Regueira, T. A., & Tomasiello, D. B. (2025). Identifying slum areas: A multidimensional analysis leveraging with explanatory machine learning techniques. Sustainable Cities and Society, 106645.

Eichholtz, P., Korevaar, M., Lindenthal, T., & Tallec, R. (2021). The total return and risk to residential real estate. The Review of Financial Studies, 34(8), 3608-3646.

Farkas, I. D. (2023). Return on investing in residential real estate Vilniaus universitetas.].

Figlioli, B., & Lima, F. G. (2022). A proposed corporate distress and recovery prediction score based on financial and economic components. Expert Systems with Applications, 197, 116726.

Gattis, L. (2023). Should I buy a Caribbean condo during a pandemic? Valuing a real estate investment using cash flows. The CASE Journal, 19(2), 137-165.

Gigante, G., & Cozzio, G. (2021). Equity crowdfunding: An empirical investigation of success factors in real estate crowdfunding. Journal of Property Investment & Finance, 40(6), 532-547.

Gopal, S., & Pitts, J. (2025). Regulation and Frameworks: Current and Future Reporting Trends. In The FinTech Revolution: Bridging Geospatial Data Science, AI, and Sustainability (pp. 183-224). Springer.

Gradillas, M., Castro, A., & Salandra, R. (2021). Understanding the adoption of voluntary environmental, social and governance (ESG) standards in the real estate sector.

Hełdak, M., & Kucher, L. (2024). Apartment Rental Market in Border Cities of Poland and Ukraine. Geomatics and Environmental Engineering, 18(2), 97-111.

Ifeanacho, K. P., & Egbenta, I. R. (2023). Application of income capitalization approach for emerging rental income cash flow in Enugu metropolis property market. Property Management, 41(3), 454-471.

Ibrahim, A., Almasria, N. A., Alhatabat, Z. A., Ershaid, D. J. A., & Aldboush, H. H. (2025). Transforming financial services with artificial intelligence and machine learning: Insights into opportunities and emerging trends. In Utilizing AI and machine learning in financial analysis (pp. 129–148).

Ibrahim, A., Almasria, N. A., Almaqtari, F. A., Al-Kasasbeh, O., Alhatabatb, Z., & Ershaid, D. (2024). The Impact of Green Finance, FinTech and Digital Economy on Environmental Sustainability: Evidence from Advanced Panel Techniques. International Journal of Energy Economics and Policy, 14(6), 621-627.

Jalihal Sharanappa, D. M. N. (2022). Quarterly Earnings Announcements And Stock Returns. Journal of Pharmaceutical Negative Results, 10925-10935.

Kasim, I., Amidu, A.-R., & Levy, D. (2025). Factors influencing property valuation models development and application as decision support systems. Property Management.

Kinatta, M. M., Kaawaase, T. K., Munene, J. C., Nkote, I., & Nkundabanyanga, S. K. (2021). Cognitive bias, intuitive attributes and investment decision quality in commercial real estate in Uganda. Journal of Property Investment & Finance, 40(2), 197-219.

Livingstone, N., & Sanderson, D. (2022). All grown up? Market maturity and investment in London's purpose-built student accommodation sector. Journal of Property Investment & Finance, 40(6), 571-587.

Madsen, M. D., Paasch, J. M., & Sørensen, E. M. (2022). The many faces of condominiums and various management structures− The Danish case. Land Use Policy, 120, 106273.

Manzano, A. (2022). Conducting focus groups in realist evaluation. Evaluation, 28(4), 406-425.

Marçal, K. E., Fowler, P. J., & Hovmand, P. S. (2023). Feedback dynamics of the low-income rental housing market: exploring policy responses to COVID-19. System Dynamics Review, 39(4), 371-403.

McGaffin, R., Viruly, F., & Boyle, L. (2019). An investigation into the use of land-based financing to fund infrastructure in South Africa. Journal of Property Investment & Finance, 39(3), 183-198.

McGough, T., & Berry, J. (2020). Pricing risk and its use in modelling real estate market yields. Journal of Property Investment & Finance, 38(5), 419-433.

Menezes, L. R. d., Lacerda, N., & Primavera, L. L. (2023). Low-income rental and heritage preservation: the case of Boa Vista (Recife). Cadernos Metrópole, 25, 1171-1199.

Mukumbang, F. C. (2023). Retroductive theorizing: a contribution of critical realism to mixed methods research. Journal of Mixed Methods Research, 17(1), 93-114.

Mustofa, M. T. L., Pratama, A. A. N., Wicaksono, P. N., Pasagi, Y., & Rini, P. (2023). ANALYSIS OF NON-PERFORMING FINANCING IN MEDIATING THE INFLUENCE OF INCOME DIVERSIFICATION, INFLATION, GROSS DOMESTIC PRODUCT, AND SIZE OF THE BANK ON STABILITY. Jurnal Ekonomi, 12(04), 933-939.

Ngoc, N. M., Tien, N. H., & Hieu, V. M. (2023). The relevance of factors affecting real estate investment decisions for post pandemic time. International journal of business and globalisation.

Ntabanganyimana, J., Dar, A. A., Nain, M., & Ahmad, M. (2025). Different Portfolio on Investment: A Practical Work to Reduce Risks. In Unveiling Investor Biases That Shape Market Dynamics (pp. 69-86). IGI Global Scientific Publishing.

Odimegwu, C., & Anyakora, M. (2023). Land value capture through land based taxation: For sustainable infrastructure finance in Anambra State, Nigeria. coou African Journal of Environmental Research, 4(1), 121-138.

Ogunbiyi, J. O. (2023). Practical Demonstrations of the Use of Select Property Investment Viability Appraisal Models: Net Present Value, Profitability Index and Net Terminal Value. Profitability Index and Net Terminal Value (May 25, 2023).

Phillips, P. P. (2023). Return on investment (ROI) basics. Association for Talent Development.

Pisal, P., Reddy, K. K., Kishore, J., Jonnalagadda, R. R., Kumar, M., Band, G., & Joshi, B. (2025). An integrated TOPSIS and ARAS method multi-criteria decision-making approach for optimizing investment portfolios using goal programming and genetic algorithm model. Scientific Reports, 15(1), 34450.

Potrawa, T., & Tetereva, A. (2022). How much is the view from the window worth? Machine learning-driven hedonic pricing model of the real estate market. Journal of Business Research, 144, 50-65.

Qiu, Z. (2023). Factors positively affecting decision to rent the condominium of Chinese tenants at rama 9 street, Huai Khwang district, Bangkok, Thailand.

Riddiough, T. J. (2022). Pension funds and private equity real estate: history, performance, pathologies, risks. In Handbook of real estate and macroeconomics (pp. 371-412). Edward Elgar Publishing.

Robin, E. (2022). Performing real estate value (s): real estate developers, systems of expertise and the production of space. Geoforum, 134, 205-215.

Robinson, D. T., & Sensoy, B. A. (2016). Cyclicality, performance measurement, and cash flow liquidity in private equity. Journal of Financial Economics, 122(3), 521-543.

Saengchote, K., & Charoenpanich, C. (2021). Cash flow uncertainty and IPO underpricing: evidence from income guarantee in Thai REITs. Journal of Property Investment & Finance, 39(6), 590-608.

Steininger, B. I., Groth, M., & Weber, B. L. (2021). Cost overruns and delays in infrastructure projects: the case of Stuttgart 21. Journal of Property Investment & Finance, 39(3), 256-282.

Szilágyi, L., Lefkovits, L., & Iclanzan, D. (2020). A review on suppressed fuzzy c-means clustering models. Acta Universitatis Sapientiae Informatica, 12(2), 302-324.

Tamilmathi, T., & Priya, R. S. (2023). Real Estate Investors Perception towards Investments. Journal for ReAttach Therapy and Developmental Diversities, 6(10s (2)), 2298-2302.

Tan, Q. C., Yijun & Wang, Xingyan & Zeng, Zexin. (2022). Studies on the Modifications and Applications of the Net Present Value and Internal Rate of Return. https://doi.org/10.2991/978-2-494069-31-2_117

Tang, B. s., Ho, W. K., & Wong, S. W. (2021). Sustainable development scale of housing estates: An economic assessment using machine learning approach. Sustainable Development, 29(4), 708-718.

Torkian, V., Bamdad, S., & Sarfaraz, A. H. (2025). Integrating AI and OR for investment decision-making in emerging digital lending businesses: a risk-return multi-objective optimization approach. Journal of the Operational Research Society, 1-20.

Walia, S., Sarkar, S., Mohanty, B., & Pal, S. (2023). Practice Briefing Analysis of the emergence and initial performance of REITs in India. Journal of Property Investment & Finance, 41(2), 256-268.

Wang, C., Chen, Y., Zhang, S., & Zhang, Q. (2022). Stock market index prediction using deep Transformer model. Expert Systems with Applications, 208, 118128.

Weber, R. (2021). Embedding futurity in urban governance: Redevelopment schemes and the time value of money. Environment and Planning A: Economy and Space, 53(3), 503-524.

Wong, W. C., Soon, J.-J., Abdullah, N. A. H., & Abdullah, M. R. B. (2025). The impact of developers' market power on the launch prices of new houses in Malaysia. International Journal of Housing Markets and Analysis.

Zhang, D., & Zhang, F. (2025). AI-Assisted Identification and Equity Assessment of Vulnerable Population Impacts in US Energy Transition. Journal of Advanced Computing Systems, 5(7), 1-17.

Zheng, Y., Zhang, J., Shukla, K. N., O'Leary, M., Wang, D. X., & Xu, J. (2025). Leveraging interactive visualizations and LLM-driven explanations for transparent multi-objective portfolio management. Discover Data, 3(1), 50.

Zhou, J. (2023). The private sector in the rental housing market. BCP Business & Management, 43, 125–129. https://api.semanticscholar.org/CorpusID:257751783.

Nicholas, T., & Scherbina, A. (2013). Real estate prices during the roaring twenties and the great depression. Real Estate Economics, 41(2), 278-309.