Geospatial Analysis of Virtual Property Prices Distributions and Clustering

Dwi Sugianto[1,*], Andhika Rafi Hananto[2]

[1]Magister of Computer Science, Universitas Amikom Purwokerto, Purwokerto, Indonesia

[2]Magister of Computer Science, Universitas Gadjah Mada, Yogyakarta, Indonesia

ABSTRACT

This paper presents an analysis of property prices in the virtual world, focusing on geographical distribution and district comparisons. Utilizing a dataset of virtual properties, we applied scatter plot analysis, cluster analysis using DBSCAN, and box plot comparison to identify key patterns and opportunities within this market. The scatter plot analysis revealed that property prices are unevenly distributed, with higher prices clustering in specific regions, indicating areas of higher desirability and value. The DBSCAN clustering identified distinct high-value clusters, each containing 10 to 67 properties, and highlighted 1,067 properties as noise, suggesting a dispersed distribution of lower-value properties. Box plot comparisons across districts showed significant variations in property values. Some districts exhibited higher median prices, with the highest at 35,452.60 MANA, while others had lower medians. Variability within districts varied, with some showing a wide range of prices and others more uniform values. Outliers suggested unique investment opportunities in both premium and undervalued properties. For virtual real estate investors, the findings emphasize the importance of location and strategic investment. High-value districts and emerging areas offer potential for significant returns. Developers and urban planners can use these insights to focus on high-demand areas, enhancing project value through strategic investments in infrastructure and amenities. This study highlights the dynamic nature of the virtual real estate market and the importance of ongoing research to understand factors influencing property values. Stakeholders can make informed decisions and capitalize on opportunities in this evolving market.

Keywords Virtual Real Estate, Property Price Analysis, Geographical Distribution, DBSCAN Clustering, Investment Opportunities

INTRODUCTION

The advent of virtual worlds has brought about significant changes in how we interact, socialize, and conduct business. As these digital environments continue to grow in complexity and popularity, they have also given rise to a new frontier in real estate investment: virtual real estate. In these virtual worlds, users can buy, sell, and trade digital properties, much like in the physical world. Understanding the dynamics of property prices in these virtual landscapes is crucial for investors, developers, and urban planners seeking to capitalize on the opportunities within this emerging market.

Virtual real estate markets share many similarities with their physical counterparts, including the influence of location on property values. Factors such as proximity to virtual landmarks, accessibility, and the presence of social or commercial hubs can significantly impact property prices [1], [2]. However, the virtual nature of these environments also introduces unique elements that can affect market dynamics, such as the availability of digital amenities and the community engagement within virtual neighborhoods [3].

Submitted 15 June 2024
Accepted 1 August 2024
Published 1 September 2024

Corresponding author
Dwi Sugianto,
dwisugianto@outlook.com

Additional Information and Declarations can be found on page 139

DOI: 10.47738/ijrm.v1i2.10

How to cite this article: D. Sugianto and A. R. Hananto, "Geospatial Analysis of Virtual Property Prices Distributions and Clustering," *Int. J. Res. Metav.*, vol. 1, no. 2, pp. 127-141, 2024.

This paper aims to analyze the geographical distribution of property prices in the virtual world and compare property values across different districts. By utilizing a dataset of virtual properties, we employ scatter plot analysis, cluster analysis using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm, and box plot comparison to identify key patterns, trends, and opportunities within this market [4], [5], [6].

The scatter plot analysis provides a visual representation of property price distribution based on geographical coordinates, highlighting areas of high demand and value [7]. The DBSCAN clustering analysis helps identify regions with high concentrations of property prices, revealing potential investment hotspots [8]. Finally, the box plot comparison across districts offers insights into the variability of property values, uncovering high-value areas and potential opportunities for investors [9].

The insights gained from this analysis are valuable for various stakeholders. Investors can use this information to make informed decisions about where to allocate their resources and focus their investment efforts. Developers and urban planners can leverage these insights to enhance the attractiveness and value of their projects by strategically investing in high-demand areas and improving infrastructure and amenities [10].

Literature Review

The virtual real estate market is a burgeoning field that parallels the dynamics of traditional real estate markets while introducing unique challenges and opportunities due to its digital nature. This literature review explores recent studies on the determinants of property prices, the application of advanced analytical techniques, and the implications of these findings for stakeholders in the virtual real estate market.

Several studies have examined the factors influencing property prices in both physical and virtual real estate markets. According to He et al. (2021), factors such as cost, interior decoration, location, and status are significant in determining real estate prices [11]. This is echoed in the findings of Chernyshova et al. (2019), who identified the influence of social, economic, and physical factors on property prices, highlighting the complex interplay of these determinants [12]. In the context of virtual real estate, Hu (2023) noted that price clustering behavior is influenced by factors such as pricing uncertainty and the number of buyers and sellers in the market [13].

The use of advanced analytical techniques has become increasingly important in understanding and predicting property prices. Abidoye et al. (2019) demonstrated that artificial intelligence (AI) techniques, such as artificial neural networks (ANN), outperform traditional models like ARIMA in predicting property price indices, providing more reliable forecasts [14]. Similarly, Ho et al. (2021) found that machine learning algorithms, particularly Random Forest (RF) and Gradient Boosting Machine (GBM), significantly enhance the accuracy of property price predictions compared to traditional methods [15].

Pashkevych et al. (2022) applied machine learning models to predict real estate prices, focusing on various regression techniques. Their analysis showed that linear regression, support vector regression, and decision tree regression are

effective in forecasting property prices, making these tools valuable for investors and developers [16].

The unique characteristics of the virtual real estate market have also been explored in several studies. Hu (2023) analyzed price clustering behavior in virtual real estate markets, documenting significant clustering at round numbers, which is consistent with the negotiation hypothesis. This behavior is influenced by pricing uncertainty and the number of buyers and sellers in the market, providing insights into the psychological factors affecting virtual real estate transactions [17].

Method

This section outlines the methodologies employed to analyze the geographical distribution of property prices in the virtual world and compare property values across different districts. The approach involves data collection, preprocessing, and the application of various analytical techniques, including scatter plot analysis, cluster analysis using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm, and box plot comparisons.

Data Collection

The dataset used for this study was obtained from a reputable virtual real estate platform, providing a comprehensive view of virtual property transactions. The dataset includes various attributes such as sales price, geographical coordinates (latitude and longitude), property type, and additional contextual information relevant to each property. These attributes are crucial for understanding the spatial distribution and value of properties in the virtual world. By incorporating detailed transaction data, the dataset allows for an in-depth analysis of property prices, helping to identify trends, patterns, and anomalies within the virtual real estate market [18].

In addition to basic transactional data, the dataset encompasses supplementary information that enriches the analysis. This includes metadata on property features, such as size and type (e.g., residential, commercial), as well as the status and history of each transaction. Contextual information, such as proximity to virtual landmarks and amenities, further enhances the dataset's utility, enabling a more nuanced examination of factors influencing property prices. The comprehensive nature of this dataset provides a robust foundation for conducting detailed analyses, facilitating a thorough exploration of the dynamics within the virtual real estate market. By leveraging such a rich dataset, this study aims to generate actionable insights for investors, developers, and urban planners operating in the digital domain [19].

Data Preprocessing

Before conducting the analysis, the dataset underwent several preprocessing steps to ensure its accuracy and consistency. The first step involved data cleaning, which is critical for maintaining the integrity of the analysis. This process included identifying and removing erroneous entries, such as duplicates or records with implausible values. Incomplete entries were addressed by handling missing values through imputation or exclusion, depending on the extent of the missing data. Additionally, discrepancies within the dataset, such as inconsistent formatting or out-of-range values, were corrected to align with standard conventions. This meticulous data cleaning

ensured that the dataset was robust and reliable, providing a solid foundation for subsequent analyses [20].

Following data cleaning, normalization was applied to the geographical coordinates (latitude and longitude) to standardize the spatial data. Normalization is essential for facilitating accurate spatial analysis, as it ensures that the coordinates are on a consistent scale, allowing for precise mapping and comparison of property locations. Furthermore, properties were categorized based on type (e.g., residential, commercial) and district, enabling meaningful comparisons across different segments of the virtual real estate market. This categorization allowed for a more granular analysis of property prices, highlighting variations and trends within specific areas and types of properties. These preprocessing steps were crucial in preparing the dataset for detailed analysis, ensuring that the findings would be both valid and insightful [21].

Scatter Plot Analysis

Scatter plot analysis was employed to visualize the distribution of property prices based on geographical coordinates, providing a clear and intuitive representation of spatial data. Each property was plotted on a two-dimensional plane, with latitude and longitude as the axes, and a color gradient used to represent the sales prices. This approach allows for an immediate visual comparison of property values across different locations. Python's Matplotlib library was utilized to create these scatter plots, offering robust features for interactive and detailed visualizations. By using color gradients to differentiate between price levels, the scatter plots effectively highlight variations in property values, making it easier to identify trends and anomalies in the virtual real estate market [22].

The visual representation provided by scatter plot analysis is invaluable for identifying regions with higher property values and potential hotspots for investment. Areas with densely clustered high-value properties stand out, indicating regions of high demand and significant market activity. Conversely, areas with lower property values or sparse distribution of high-priced properties can also be easily identified. This spatial insight is crucial for investors looking to maximize returns by targeting high-demand areas and for urban planners aiming to understand the dynamics of property distribution. The scatter plot not only provides immediate visual insights into the spatial distribution of property prices but also serves as a foundational tool for further spatial analysis, such as identifying the impact of proximity to virtual landmarks and infrastructure on property values [23].

Cluster Analysis using DBSCAN

The DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm was applied to identify clusters of high-value properties and regions with significant concentrations of property prices. DBSCAN is particularly well-suited for this analysis as it effectively handles data with noise and varying densities, making it ideal for the heterogeneous landscape of virtual real estate. The algorithm works by identifying densely packed groups of data points and distinguishing them from outliers or noise. For our analysis, the key parameters, eps (the maximum distance between two samples for them to be considered as part of the same neighborhood) and min_samples (the minimum number of samples in a neighborhood for a point to be considered a core point), were

carefully optimized to accurately reflect the natural clustering behavior present in the dataset. This optimization ensures that the clusters formed are meaningful and representative of actual high-value regions within the virtual real estate market [24].

The clustering was implemented using the Scikit-learn library in Python, which provides robust tools for machine learning and data analysis. The results of the DBSCAN algorithm categorized properties into distinct clusters, highlighting areas with dense concentrations of high-value properties. Properties classified as noise were those that did not belong to any significant cluster, indicating a more dispersed distribution of values. This clustering analysis is crucial for understanding the spatial dynamics of property prices, as it reveals high-demand areas that are prime for investment. By identifying these clusters, investors can target regions with the highest potential for return on investment, while developers and urban planners can focus their efforts on enhancing the infrastructure and appeal of these high-value zones. The insights gained from the DBSCAN clustering provide a detailed map of property value distribution, guiding strategic decision-making in the virtual real estate market [25].

Box Plot Comparison

Box plot analysis was employed to compare property prices across different districts, providing a clear and concise visualization of price distributions within each district. Box plots are particularly effective for this purpose as they summarize the central tendency and variability of data, including median prices, interquartile ranges (IQR), and potential outliers. In this study, the Seaborn library in Python was utilized to generate these box plots, leveraging its advanced plotting capabilities and detailed customization options. The visualizations produced by Seaborn allow for an intuitive comparison of property values across various districts, making it easier to identify trends and anomalies within the dataset. This methodological choice ensures that the analysis is both accurate and accessible, providing stakeholders with valuable insights into the virtual real estate market [26].

The box plots revealed significant variations in property prices across different districts. Some districts exhibited notably higher median prices, indicating areas of greater desirability and value. The interquartile range in each box plot highlighted the spread of property prices within the districts, with some districts showing a wide range of values and others demonstrating more uniform pricing. The presence of outliers in certain districts suggested unique investment opportunities, either in high-value premium properties or undervalued assets with potential for appreciation. These outliers are critical for investors seeking to identify properties that deviate significantly from the median, offering prospects for high returns. By visualizing these differences, the box plot analysis aids in pinpointing districts that warrant further investigation and investment, helping investors and developers make informed decisions based on the distribution of property prices [27].

Result and Discussion

Results

The scatter plot analysis of property prices based on geographical coordinates (latitude and longitude) in the virtual world offers a detailed visual representation

of the spatial distribution of property values. In Figure 1, each point represents an individual property, with the color gradient indicating the sales price, ranging from lower to higher values. This method of visualization is instrumental in understanding how property prices vary across different regions within the virtual environment.

Upon examining the scatter plot, we observe that property prices are not evenly distributed across the virtual landscape. Instead, there are specific regions where property prices are noticeably higher, suggesting that these areas possess characteristics that make them more desirable or valuable. These characteristics might include proximity to virtual landmarks, accessibility to amenities, or the presence of popular social or commercial hubs. The clustering of high-priced properties in these areas indicates potential hotspots for property investments, where demand is likely to be higher.

One significant observation from the scatter plot is the concentration of high-priced properties within certain latitude and longitude ranges. These clusters of high value are indicative of areas where the virtual real estate market is particularly active and competitive. For instance, properties located near central plazas or popular districts tend to have higher prices due to the increased foot traffic and visibility, much like prime real estate in the physical world.

Furthermore, the scatter plot reveals several trends and patterns that are crucial for understanding the virtual property market. For example, properties located along major virtual roads or near transportation hubs often show higher values, reflecting their importance for accessibility and connectivity. Similarly, properties within close proximity to significant landmarks or attractions tend to be priced higher, highlighting the value of strategic location within the virtual world.

The visual representation provided by the scatter plot is not only useful for identifying high-value areas but also for spotting potential investment opportunities. Areas that show a gradual increase in property prices might be indicative of emerging hotspots, where early investment could lead to substantial returns. Conversely, regions with consistently lower prices might suggest areas with less demand or undeveloped potential, which could be targeted for future development projects.

This scatter plot analysis also facilitates a deeper understanding of the virtual world's economic dynamics. It allows stakeholders, such as investors, developers, and urban planners, to make data-driven decisions based on the spatial distribution of property prices. By identifying regions with higher demand and value, stakeholders can allocate resources more effectively, plan strategic developments, and optimize their investment portfolios.

Additionally, the scatter plot serves as a foundational tool for further spatial analyses, such as identifying the impact of virtual infrastructure developments on property values or exploring the relationship between social engagement and real estate prices. These insights are invaluable for developing comprehensive strategies that enhance the overall value and attractiveness of the virtual world.

The scatter plot analysis of property prices based on geographical coordinates is a powerful tool for visualizing and understanding the distribution of property values in the virtual world. It highlights regions of high demand and value, identifies potential investment hotspots, and provides a basis for strategic

decision-making. The detailed patterns and trends observed in the scatter plot are essential for stakeholders aiming to navigate the virtual real estate market effectively and capitalize on emerging opportunities as seen in figure 1.

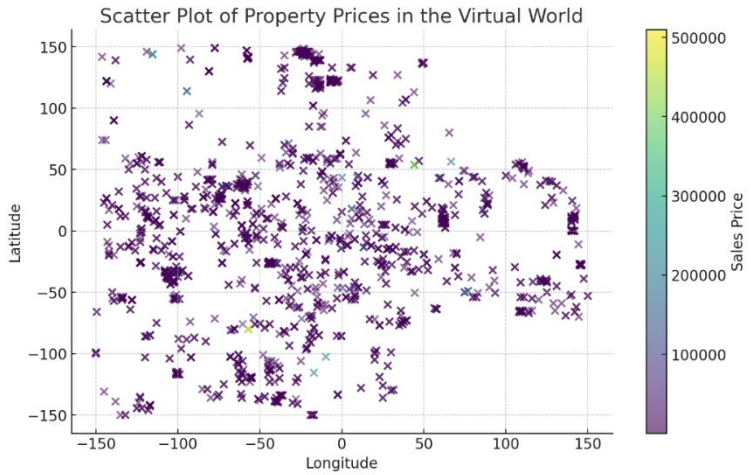

Figure 1 Scatter plot of property prices based on geographical coordinates

Cluster Analysis

To further delve into the geographical distribution of property prices in the virtual world, we conducted a clustering analysis utilizing the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. DBSCAN is particularly well-suited for identifying clusters within datasets characterized by noise and varying densities, making it an ideal choice for our analysis of virtual property prices.

The DBSCAN algorithm operates by grouping together points that are closely packed together, marking as outliers points that lie alone in low-density regions. This method is highly effective in distinguishing between densely populated clusters and dispersed noise, which is crucial for understanding the spatial distribution of property prices.

Upon applying DBSCAN to our dataset, we found that the majority of properties (1,067) were classified as noise, labeled as cluster -1. This classification indicates that these properties do not belong to any significant cluster, suggesting a widely dispersed distribution of property prices across the virtual landscape. The presence of such a large number of noise points reflects a broad spread in property values without clear, dense clusters in many areas of the virtual world.

However, the DBSCAN analysis also revealed several distinct clusters, each containing between 10 to 67 properties. These clusters represent regions with high concentrations of property prices, indicating areas of elevated demand or value. The identification of these clusters is pivotal for understanding where the most valuable properties are located and the spatial dynamics of the virtual real estate market.

Cluster Analysis Insights:

The clusters identified by DBSCAN highlight regions where property prices are significantly higher than in the surrounding areas. These high-value areas are

likely to be hotspots for virtual real estate investment, offering opportunities for substantial returns due to their desirability.

By analyzing the spatial distribution of these clusters (figure 2), we can gain insights into the factors driving property values in the virtual world. For instance, clusters may form around virtual landmarks, popular social hubs, or regions with superior virtual infrastructure. Understanding these dynamics helps in pinpointing why certain areas command higher prices.

The clusters provide valuable information for potential investors. Properties within these high-value clusters are prime candidates for investment due to their higher demand and potential for appreciation. Conversely, understanding areas classified as noise can help investors identify regions that may be undervalued or lack development, presenting opportunities for future growth.

For developers and urban planners in the virtual world, the identification of clusters aids in strategic planning. High-value clusters indicate where to focus development efforts to maximize property values and attract more users. Additionally, recognizing areas of noise can highlight regions that might benefit from infrastructure improvements or new attractions to boost their value.

The clustering analysis also facilitates market segmentation by dividing the virtual world into high-value clusters and dispersed regions. This segmentation is useful for tailoring marketing strategies, setting property prices, and understanding the different market dynamics at play in various regions of the virtual world.

The DBSCAN clustering not only helps in identifying high-value areas but also provides a comprehensive understanding of the spatial structure of the virtual property market. By visualizing these clusters, stakeholders can make informed decisions regarding virtual property investments, development, and marketing strategies.

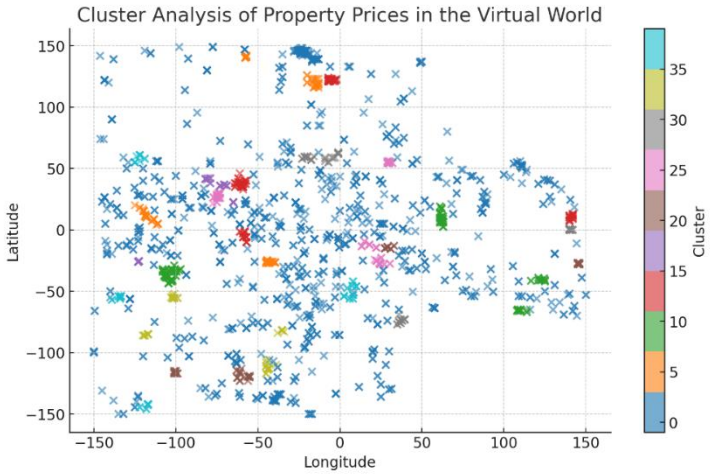

Figure 2 Cluster analysis of property prices in the virtual world

Comparison of Property Prices by District

To gain a comprehensive understanding of property prices across different districts in the virtual world, we employed a box plot analysis. This method allows us to visualize the distribution of prices within each district, providing

insights into the central tendency and variability of property values. By examining the median prices, interquartile ranges (IQR), and outliers, we can identify significant patterns and disparities in property prices across districts (figure 3).

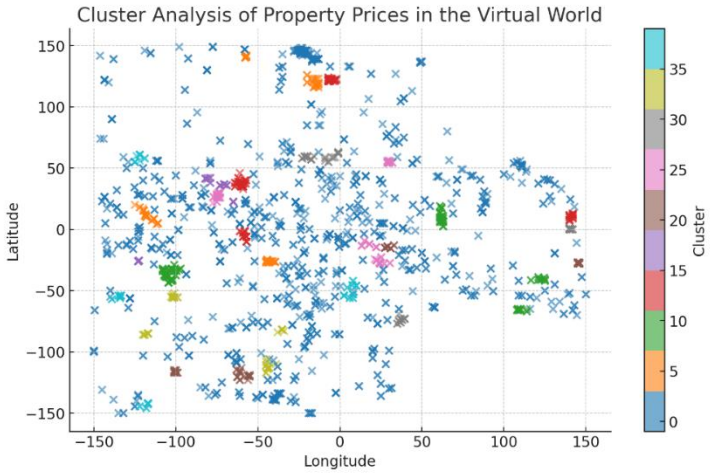

Figure 3 Comparison of property prices by district

The box plot reveals the median property prices for each district, represented by the line within each box. Districts with higher median prices indicate areas of greater value and demand. These districts are likely to be more developed or have attributes that make them particularly desirable, such as proximity to virtual landmarks, superior infrastructure, or vibrant social and commercial activities. Higher median prices in certain districts reflect the premium that users are willing to pay for properties in these sought-after locations.

The interquartile range, represented by the height of each box, shows the middle 50% of property prices within each district. A wider IQR indicates greater variability in property prices, suggesting a mix of high and low-value properties within the district. This could be due to diverse property types, varying levels of development, or different levels of accessibility and amenities. Conversely, a narrower IQR suggests more uniform property prices, indicating a more homogenous district in terms of property value and possibly development status.

The presence of outliers, represented by individual points outside the whiskers of the box plot, highlights properties with exceptionally high or low prices compared to the majority within the same district. High outliers may indicate premium properties with unique attributes, such as prime locations, larger sizes, or special features. Low outliers, on the other hand, may represent undervalued properties or those requiring development. The identification of outliers is crucial for investors, as these properties may offer unique opportunities for investment, either as high-end premium assets or as undervalued properties with potential for appreciation.

The box plot analysis highlights significant variations in property prices across different districts. Some districts exhibit notably higher median prices and a wide range of values, indicating areas of mixed development and diverse property offerings. These districts may attract a variety of investors and users, from those seeking premium properties to those looking for more affordable options. Other

districts show more uniform prices, suggesting a stable and consistent property market within those areas. Understanding these variations is essential for investors to identify high-value areas and tailor their investment strategies accordingly.

Districts with higher median prices and fewer outliers are indicative of high-value areas with strong and stable property markets. These districts are likely to offer better returns on investment due to consistent demand and higher property values. Investors can prioritize these districts for secure and profitable investments. Additionally, these high-value districts may serve as benchmarks for development in other areas, setting standards for infrastructure, amenities, and overall attractiveness.

Districts with a wide IQR and numerous outliers, particularly on the higher end, may represent emerging areas with significant potential for growth and development. These districts could be undergoing rapid changes, attracting new investments and developments that drive up property values. Identifying these emerging districts early can provide investors with opportunities to capitalize on future growth and appreciation in property prices.

The box plot comparison of property prices by district provides valuable insights for virtual real estate investors. By understanding the distribution of property values within each district, investors can make informed decisions about where to allocate their resources and focus their investment efforts. High-value districts with stable property prices offer secure investment opportunities, while emerging districts with varied property values present opportunities for high returns through strategic investments and developments.

Investors can use this analysis to identify districts that align with their investment goals, whether seeking stable, high-value properties or targeting areas with potential for significant growth. Additionally, the presence of outliers highlights unique investment opportunities that may offer substantial returns, either through premium properties or undervalued assets with development potential.

For developers and urban planners, the box plot analysis provides insights into the current state of property markets across different districts. Understanding the variability and distribution of property prices helps in planning future developments, optimizing resource allocation, and enhancing the overall attractiveness of different areas. By focusing on high-value districts and emerging areas, developers can create targeted strategies that align with market demand and maximize the potential for growth and profitability.

In conclusion, the comparison of property prices by district through box plot analysis offers a detailed and nuanced understanding of the virtual real estate market. This analysis highlights significant variations in property values, identifies high-value and emerging districts, and provides essential insights for investors, developers, and urban planners. By leveraging this information, stakeholders can make informed decisions, optimize their strategies, and capitalize on opportunities within the dynamic virtual real estate market.

Discussion

The geographical distribution of property prices in the virtual world is not uniform. The scatter plot analysis reveals that certain regions exhibit higher property prices, indicating areas of higher demand or value. These regions are

likely to have attributes that make them more attractive, such as proximity to virtual landmarks, accessibility, and the presence of popular activities or services. The clustering of high-priced properties in specific areas suggests the existence of virtual "neighborhoods" with varying levels of desirability. Similar to real-world real estate markets, factors such as location, accessibility, and amenities play crucial roles in determining property values in the virtual world.

The cluster analysis using DBSCAN identified several high-concentration price areas. These clusters are regions where property prices are significantly higher than in the surrounding areas. The presence of these clusters indicates that certain locations within the virtual world are more valuable due to factors such as virtual infrastructure, community engagement, or strategic importance. Understanding these high-concentration areas can provide valuable insights for virtual real estate investors. Investing in properties within these clusters could yield higher returns due to the increased demand and value associated with these locations. Additionally, recognizing the characteristics of these clusters can help investors make informed decisions about future investments and development in the virtual world.

The comparison of property prices by district reveals significant differences in property values across various districts. Some districts have higher median prices, indicating they are more desirable or valuable locations. Factors contributing to these variations could include the quality of virtual infrastructure, the presence of key landmarks or attractions, and the level of community engagement within each district. Investors can use this information to identify high-value districts and prioritize their investments accordingly. Districts with higher median prices and fewer outliers may offer more stable investment opportunities, while districts with a wider range of prices might present opportunities for higher risk and reward.

For virtual real estate investors, the findings from this analysis provide several key insights. Firstly, location matters: similar to real-world real estate, the location of a property in the virtual world significantly impacts its value. High-demand areas and clusters of high-priced properties are likely to yield better investment returns. Secondly, cluster identification is crucial: identifying clusters of high-priced properties can help investors target specific areas for investment, maximizing their potential returns. Thirdly, understanding variations across districts allows investors to make informed decisions about where to invest based on their risk tolerance and investment goals.

Conclusion

The analysis of property prices in the virtual world, based on geographical coordinates and district comparisons, has provided significant insights into the distribution and value of virtual real estate. By employing scatter plot analysis, cluster analysis using DBSCAN, and box plot comparison across districts, we have been able to identify key patterns, trends, and opportunities within this emerging market.

The geographical distribution of property prices reveals that they are not evenly spread across the virtual landscape. The scatter plot analysis highlighted those high-priced properties tend to cluster in specific regions, suggesting that these areas possess attributes that make them more desirable and valuable. Factors such as proximity to virtual landmarks, ease of accessibility, and the presence

of social or commercial hubs are likely contributors to the higher prices in these hotspots. This insight helps stakeholders understand where demand is concentrated and where the highest values are found.

The DBSCAN clustering analysis identified several distinct clusters of high-value properties, pinpointing regions with significantly elevated property prices. These clusters represent areas with high demand, making them prime targets for investment. The large number of properties categorized as noise indicates a dispersed distribution of lower-value properties across the virtual world. This differentiation between clustered high-value areas and widely spread lower-value properties provides a clearer picture of the virtual real estate market dynamics.

Comparing property prices across different districts through box plot analysis revealed substantial variations in property values. Some districts exhibit notably higher median prices, indicating these areas are more valuable and desirable. The variability in property prices within each district varies, with some districts showing a wide range of values and others more uniform. The presence of outliers in certain districts suggests there are exceptional properties with either very high or very low prices compared to the majority within the same district. These outliers highlight unique investment opportunities, such as premium properties or undervalued assets.

For virtual real estate investors, the findings emphasize the importance of location and strategic investment. High-value districts and identified clusters offer prime opportunities for significant returns on investment. Recognizing the variability within districts and spotting emerging areas with potential for growth can help investors make informed decisions, thereby optimizing their investment portfolios. Understanding these insights is crucial for navigating the virtual real estate market effectively.

Developers and urban planners can use these insights to focus on high-value and emerging districts, enhancing the attractiveness and value of their projects. By aligning development efforts with market demand and strategically investing in infrastructure and amenities, developers can maximize growth and profitability. This strategic focus ensures that developments are positioned in areas with the highest potential for success, benefiting both the developers and the virtual communities they serve.

Further research should delve into additional factors influencing property prices, such as the impact of virtual events, levels of community engagement, and the role of infrastructure development. Longitudinal studies can provide valuable insights into how property values evolve over time, identifying trends and informing future investment and development strategies. This ongoing research is essential for understanding the dynamic nature of the virtual real estate market and staying ahead of its developments.

The comprehensive analysis of property prices in the virtual world has highlighted the significance of geographical distribution, identified high-value clusters, and revealed variations across districts. These findings offer valuable insights for investors, developers, and urban planners, guiding them in making informed decisions and capitalizing on opportunities within the dynamic virtual real estate market. By continuing to explore and understand the factors driving property values, stakeholders can navigate this emerging market effectively and

achieve sustained success.

Declarations

Author Contributions

Conceptualization: D.S., and A.R.H.; Methodology: A.R.H.; Software: D.S.; Validation: D.S.; Formal Analysis: D.S.; Investigation: D.S.; Resources: D.S.; Data Curation: D.S.; Writing Original Draft Preparation: D.S.; Writing Review and Editing: A.R.H.; Visualization: D.S. and A.R.H.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] B. Hu, "Price clustering behavior in virtual real estate markets," Applied Finance Letters, vol. 12, no. 1, pp. 1-13, 2023.

[2] G. Lisi and M. Iacobini, "Measuring the effect of location on house prices in Italy," Regional Economic Development Research, vol. 12, no. 2, pp. 255-270, 2020.

[3] X. Zhou, "VRML-based real estate 3D virtual scene display and its impact on property prices," Multimedia Tools and Applications, vol. 110, no. 4, pp. 1-13, 2019.

[4] R. Kaźmierczak, "Geoinformation support system for real estate market," Acta Scientiarum Polonorum Administratio Locorum, vol. 12, no. 3, pp. 350-367, 2020.

[5] A. Szczepańska, D. Gosciewski, and M. Gerus-Gosciewska, "A GRID-based spatial interpolation method as a tool supporting real estate market analyses," ISPRS Int. J. Geo Inf., vol. 9, no. 1, pp. 39, 2020.

[6] M. Kobylińska and R. Cellmer, "Modelling and simulation of selected real estate market spatial phenomena," ISPRS Int. J. Geo Inf., vol. 8, no. 10, pp. 446, 2019.

[7] Y. Boriskina, "Digital and online technology changing the real estate market in Russia," Real Estate: Economics, Management, vol. 10, no. 2, pp. 34-88, 2021.

[8] A. Adegoke, T. Oladokun, T. O. Ayodele, S. Agbato, and A. Jinadu, "DEMATEL method of analysing the factors influencing the decision to adopt virtual reality

technology by real estate firms in Lagos property market," Smart and Sustainable Built Environment, vol. ahead-of-print, no. 1, pp. 118-132, 2021.

[9] A. Szczepańska, "Transport accessibility in a suburban zone and its influence on the local real estate market: A case study of the Olsztyn functional urban area (Poland)," Land, vol. 10, no. 5, pp. 465, 2021.

[10] V. Koktashev, V. Makeev, P. Peresunko, A. Mikhalev, and V. Tynchenko, "Comparison of prices depending on factors in the secondary housing market," SHS Web of Conferences, vol. 11, no. 2, pp. 83, 2021.

[11] H. He, Y. Chen, J.-Y. Xiao, X. Chen, and Z. Lee, "Data Analysis on the Influencing Factors of the Real Estate Price," Artificial Intelligence Evolution, vol. 22, pp. 255-270, 2021.

[12] M. Chernyshova, A. Malenkaya, and T. Mezhuyeva, "Analysis of Pricing Factors in Real Estate Market," Interexpo GEO-Siberia, vol. 6, no. 2, pp. 79-85, 2019.

[13] B. Hu, "Price clustering behavior in virtual real estate markets," Applied Finance Letters, vol. 12, no. 1, pp. 1-13, 2023.

[14] R. Abidoye, A. Chan, F. A. Abidoye, and O. Oshodi, "Predicting property price index using artificial intelligence techniques," International Journal of Housing Markets and Analysis, 2019.

[15] W. K.O. Ho, B. Tang, and S. Wong, "Predicting property prices with machine learning algorithms," Journal of Property Research, vol. 38, no. 1, pp. 48-70, 2021.

[16] O. Pashkevych, S. Vahschyshchak, A. Boichuk, T. Styslo, and M. Demchyna, "Application of machine learning models for predicting prices on the real estate market," Herald of Khmelnytskyi National University. Technical sciences, vol. 313, no. 5, pp. 265-273, 2022.

[17] B. Hu, "Price clustering behavior in virtual real estate markets," Applied Finance Letters, vol. 12, no. 1, pp. 1-13, 2023.

[18] M. Cajias, P. Freudenreich, A. Freudenreich, and W. Schäfers, "Liquidity and prices: a cluster analysis of the German residential real estate market," Journal of Business Economics, vol. 90, pp. 1021-1056, 2020.

[19] M. Hahsler, M. Piekenbrock, and D. Doran, "dbscan: Fast Density-Based Clustering with R," Journal of Statistical Software, vol. 91, no. 1, 2019.

[20] M. Kobylińska and R. Cellmer, "Modelling and simulation of selected real estate market spatial phenomena," ISPRS Int. J. Geo Inf., vol. 8, no. 10, 2019.

[21] A. Kopnin and D. Prokoshev, "Fuzzy model for assessing the impact of pricing factors on the cost of primary housing," Digital Models and Solutions, vol. 2, pp. 2-4, 2023.

[22] V. Koktashev, V. Makeev, P. Peresunko, A. Mikhalev, and V. Tynchenko, "Comparison of prices depending on factors in the secondary housing market," SHS Web of Conferences, vol. 11, no. 2, pp. 83, 2021.

[23] H. He, Y. Chen, J.-Y. Xiao, X. Chen, and Z. Lee, "Data Analysis on the Influencing Factors of the Real Estate Price," Artificial Intelligence Evolution, vol. 22, pp. 255-270, 2021.

[24] Y. Xie and S. Shekhar, "Significant DBSCAN towards Statistically Robust Clustering," Proceedings of the 16th International Symposium on Spatial and

Temporal Databases, 2019.

[25] Z. Li, Y. Li, W. Lu, and J. Huang, "Crowdsourcing Logistics Pricing Optimization Model Based on DBSCAN Clustering Algorithm," IEEE Access, vol. 8, pp. 92615-92626, 2020.

[26] T. Boonchoo, X. Ao, Y. Liu, W. Zhao, F. Zhuang, and Q. He, "Grid-based DBSCAN: Indexing and inference," Pattern Recognit., vol. 90, pp. 271-284, 2019.

[27] E. Ozhegov and A. Ozhegova, "Distance in geographic and characteristics space for real estate pricing," International Journal of Housing Markets and Analysis, vol. 14, no. 3, 2021.