Green Spaces and Real Estate: Exploring the Influence of Environmental Factors on Housing Prices in Busan

CHEN Peng Wei, MOO Jia Rong and ZHANG Xiao Han

Singapore Management University

Introduction

Issue & problem: With rapid urbanization, the impact of neighborhood environments on well-being and property values has gained attention. Research shows that access to green spaces enhances well-being and increases housing prices. While this link is well-documented in other regions, its impact in Korea—particularly regarding proximity to greenery, environmental amenities, and the built environment—remains understudied.

Motivation: As cities rapidly urbanize, understanding how these factors influence property values is crucial for sustainable urban planning. Green spaces, environmental amenities (e.g., proximity to water bodies), and the built environment (e.g., public transit, schools) shape both livability and real estate markets, yet their integration into urban planning is often overlooked. Thus, this study aims to take a holistic approach by examining various factors influencing housing prices, with a primary focus on these three aspects.

Insights from this research can help developers optimize land use, while policymakers can strategically incorporate green spaces and infrastructure to enhance livability and sustain property values. Without data-driven planning, unchecked expansion may reduce quality of life and destabilize the real estate market.

Dataset

The dataset comprises of comprehensive information on the **property prices and characteristics** in the **Busan** Metropolitan City of South Korea for transactions in 2018 to 2019.

The original variable includes:

1. **Property Price**
2. Spatial Coordinates: Longitude, Latitude
3. **Property Characteristics:** Size, Floor, Highest Floor, Units, Parking (Number of parking spaces per household), Heating type (Dummy variable), Year of construction
4. **Environmental Amenities:** Dist. Green, Dist. Water, Green Index
5. **Local Built Environment:** Dist. Subway, Bus Stop, Dist. CBD, Top Univ., High School
6. Local Demographics: Sex Ratio, Population, Pop. Density, Higher Degree, Young Population, Median Age, Old Population
7. Sales Period (Dummy variable) : Spring, Fall, Winter

In total, there are 28 variables. Out of these, 4 are dummy variables, and the remaining 24 are numerical variables. Please note that the variables used in different analyses may vary depending on the requirements of each analysis. Binning is also applied to several continuous variables for better interpretability.

Shiny App Modules

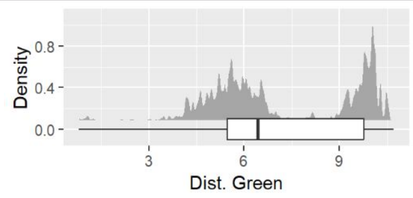
We will analyse and visualise the relationships between different factors and the housing prices in Busan by creating interactive R shiny applications. Users will also be able to analyse the data based on their interests and draw their own insights and conclusions regarding the housing prices in Busan.

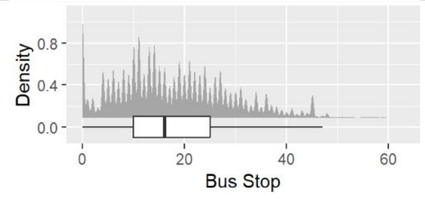
Our approach includes **4 shiny sub-modules:**

- **Exploratory Data Analysis:** Univariate and multivariate analyses
- **Confirmatory Data Analysis:** Correlation analysis and two-sample mean test/ANOVA
- **Explanatory Modeling:** Multiple linear regression and latent class analysis
- **Predictive Modeling:** Decision tree and random forest

By performing the analysis above, we can explore the relationship between price and other independent variables. The 4 sub-modules are integrated into a R shiny dashboard and is linked to our main project website.

Exploratory & Confirmatory Data Analysis



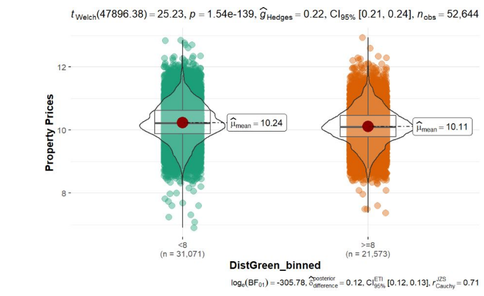


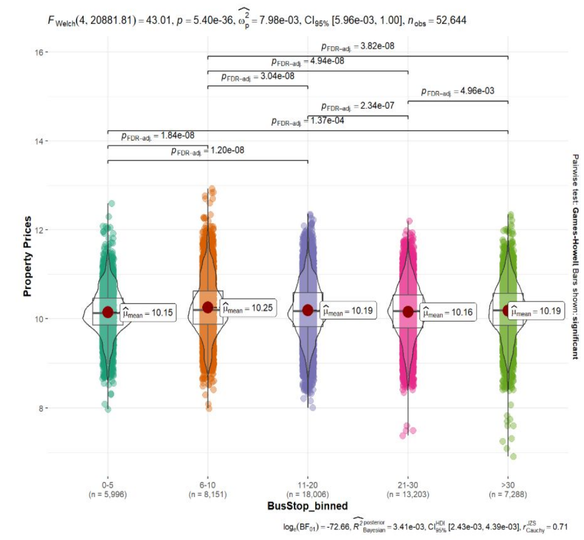
EDA:

- There are 2 main distribution groups for distance to greenery from house: <8 and ≥8
- Bus stop counts nearby house varies widely, with median number of 16 bus stops within 400m radius.

CDA:

- Houses with **6–10 nearby bus stops** have the **highest prices**, suggesting an optimal balance of accessibility and convenience.
- <6 bus stops may indicate limited accessibility to public transport, while >10 might not add value due to noise, congestion, or diminishing returns.
- Properties **closer to greenery** command significantly **higher prices.**





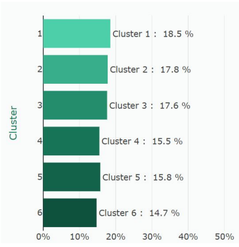
Latent Class Analysis

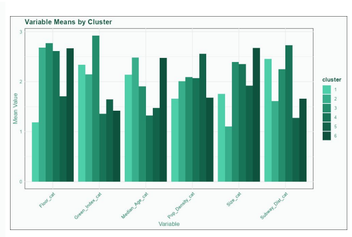
After evaluating models with 3-8 clusters, we found that **6-cluster solution** (BIC: 682,642.2) provides the **best balance between statistical fit and interpretability**. While the 7-class model showed slightly lower BIC (682,028.5), the improvement was marginal and resulted in less distinct classes.

Cluster	Proportion	Size	Floor	Green Index	Subway Access	Population Density	Key Characteristics
1	11.63%	High	High	Medium	Medium	Medium	Premium segment: Large properties on high floors appealing to affluent buyers prioritizing space and views
2	34.04%	Medium	Medium	Medium	Medium	Medium	Mainstream segment: Balanced attributes across all variables, representing typical Busan properties
3	7.00%	Medium	Medium	High	Medium	High	Urban-environmental niche: Medium-sized properties where green amenities offset urban density
4	11.14%	Medium -High	Medium	Medium-Low	Low	Medium-Low	Spacious peripheral: Larger properties where space compensates for poor subway accessibility
5	21.20%	Medium	Medium	High	Medium	Low	Green suburban: Properties in low-density areas with high environmental quality, appealing to families
6	15.00%	Low	Medium	Medium	High	High	Urban convenience: Smaller units with excellent subway accessibility where location advantages offset size limitations

Market Insights

- Environmental quality creates value across multiple segments
- Clear trade-offs between size and accessibility
- Green amenities particularly valued in Clusters 3 and 5

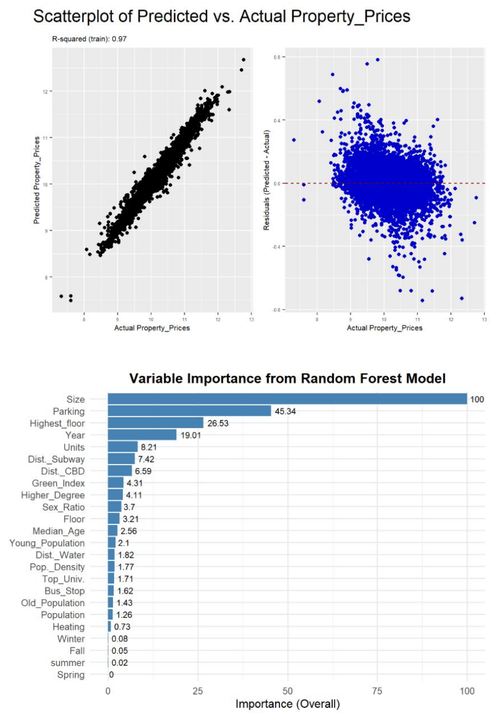




Predictive Modelling

Random Forest:

- Data Preparation: Apply a logarithmic transformation to highly skewed variables and remove the variable with the least feature importance that has high correlations (above 0.8).
- Model Evaluation: Based on the chart, the points were closely aligned with the line, confirming the model's high predictive accuracy (test R-squared of 0.972). The residuals were mostly centered around 0 with minimal spread, indicating that the model's predictions were highly accurate and lacked systematic bias.
- Variable Importance: The plot highlighted that **Size, Parking, Highest_Floor, Year, and Units** were the most influential predictors of property prices. Regression Tree also shows the similar result. Interestingly, Parking and Highest_Floor emerged as unexpected influencers, since common sense would suggest that floor level and distance to the subway have a greater influence on property prices. This might be related to Busan's unique geographic characteristics.



Conclusion & Future Directions

- This research **confirms that environmental factors**, including green spaces and public transit access, **significantly influence housing prices in Busan**. Properties near greenery commanded higher prices, and having 6–10 stops nearby is optimal for maximizing value.
- The optimal 6-cluster solution (BIC: 682,042.2) provided clear interpretability with distinct classes. Environmental quality creates value across market segments, especially in Clusters 3 and 5, enabling clear trade-offs between size and accessibility. These insights can guide developers and policymakers in optimizing land use and infrastructure to enhance livability while sustaining property values.
- **Predictive modeling showed size and parking as key price predictors**, with year of construction and highest floor also impacting values. Model evaluation demonstrated strong alignment between predictions and actual values with minimal bias.
- Future research **should explore temporal changes in environmental preferences, climate resilience factors in pricing models, and comparisons with other Asian cities** to provide deeper insights into regional differences and guide sustainable urban planning that balances livability with market stability.