Rental Price Analysis in Estonian Cities: A Geospatial Approach

Project ID H9

Understanding Spatial Factors Affecting Rental Prices in Tartu and Tallinn

Mari Lee Lumberg • Karl Martin Puna • Daniel Henri Trump

1. Introduction & Motivation

Motivation: Rental markets are spatially structured—location and accessibility shape price variation. We examine how network-based amenity distances and spatial autocorrelation influence rental prices in Tartu and Tallinn.

Goal: Quantify spatial effects on €/m² and build predictive models that account for spatial dependence.

2. Data & Sources

Rental listings: KV.ee (web-scraped)

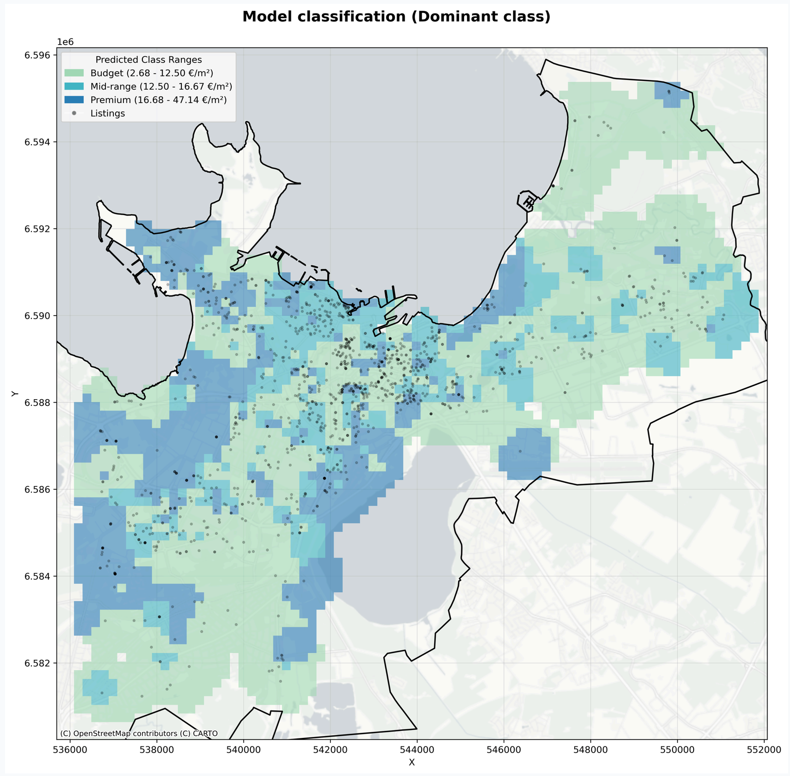
Street network & amenities: OpenStreetMap via OSMnx - amenities grouped into aggregated categories

Geospatial data of neighbourhoods and districts: Maa- ja Ruumiamet (Estonian Land and Spatial Development Board)

3. Data Cleaning

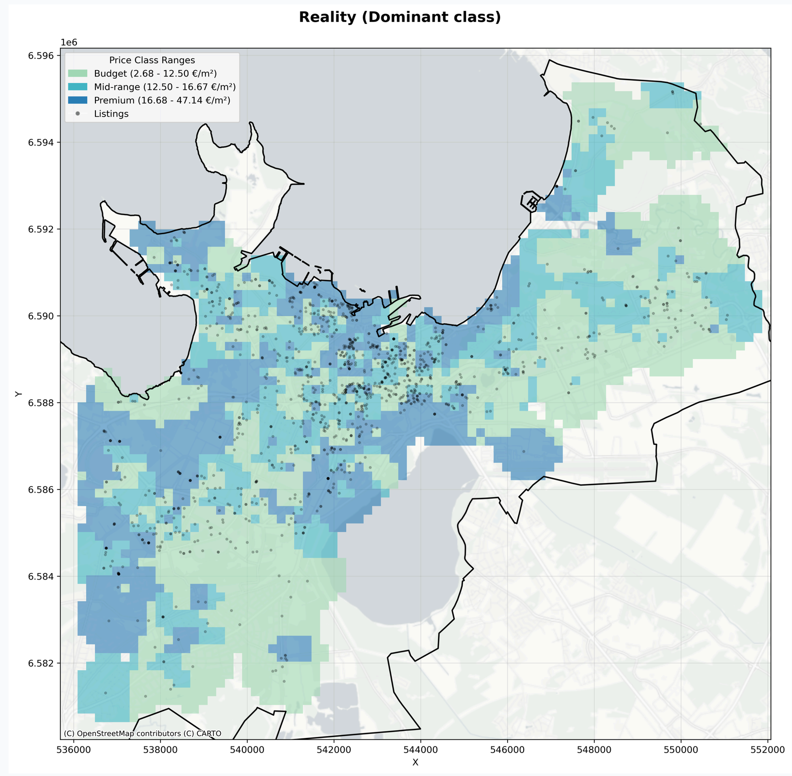
Rental listings cleaning: Tallinn (1,359), Tartu (450). Cleaned price/area, parsed floor data, removed missing/irrelevant columns (>20% missing).

OSM data cleaning: Converted polygons to centroids, filtered relevant POIs, aggregated into amenity groups



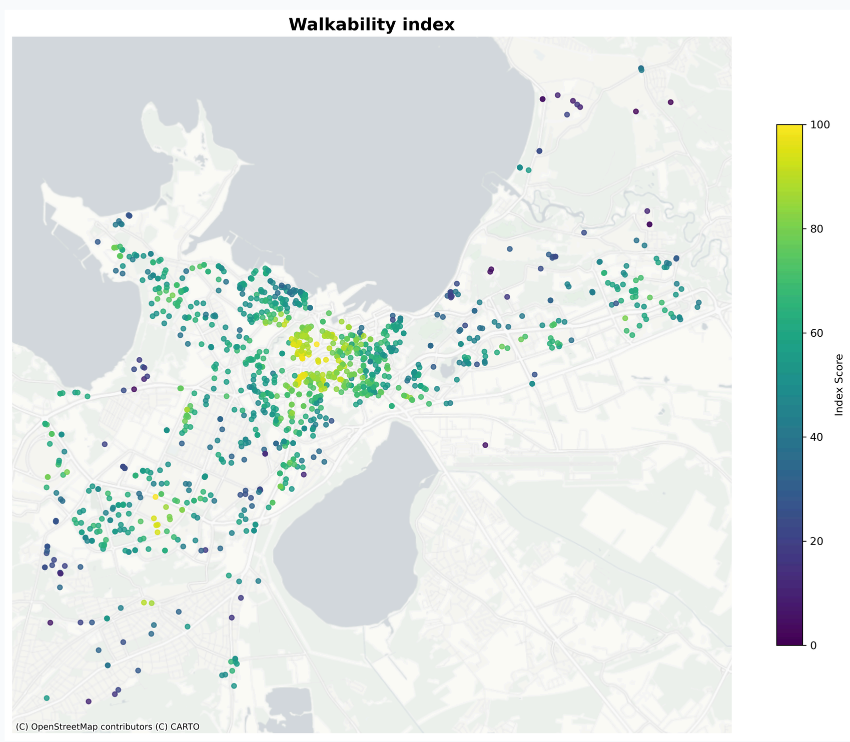
Tallinn Model Classification

Predicted price class per 200m grid (Green=Budget, Blue=Mid-range, Purple=Premium). Black dots=listings.



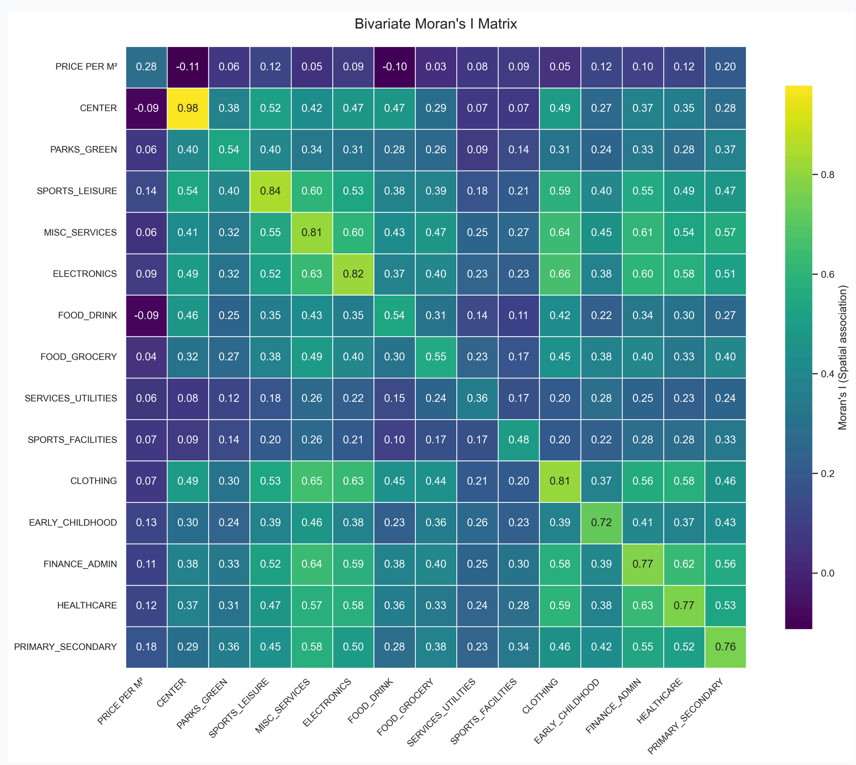
Tallinn Reality Map

Actual dominant price class based on listings, interpolated to 200m grid. Shows real spatial patterns; black dots = property listings.



Walkability Index

Composite score (0-100) based on weighted amenity proximity.



Bivariate Moran's I Matrix

Spatial correlations in Tallinn. Positive values (≥0.5): High-high matches. Negative values (≤-0.1): High-low matches. -0.1 to 0.1: Weak/no relation.

Reading the Correlation Matrix

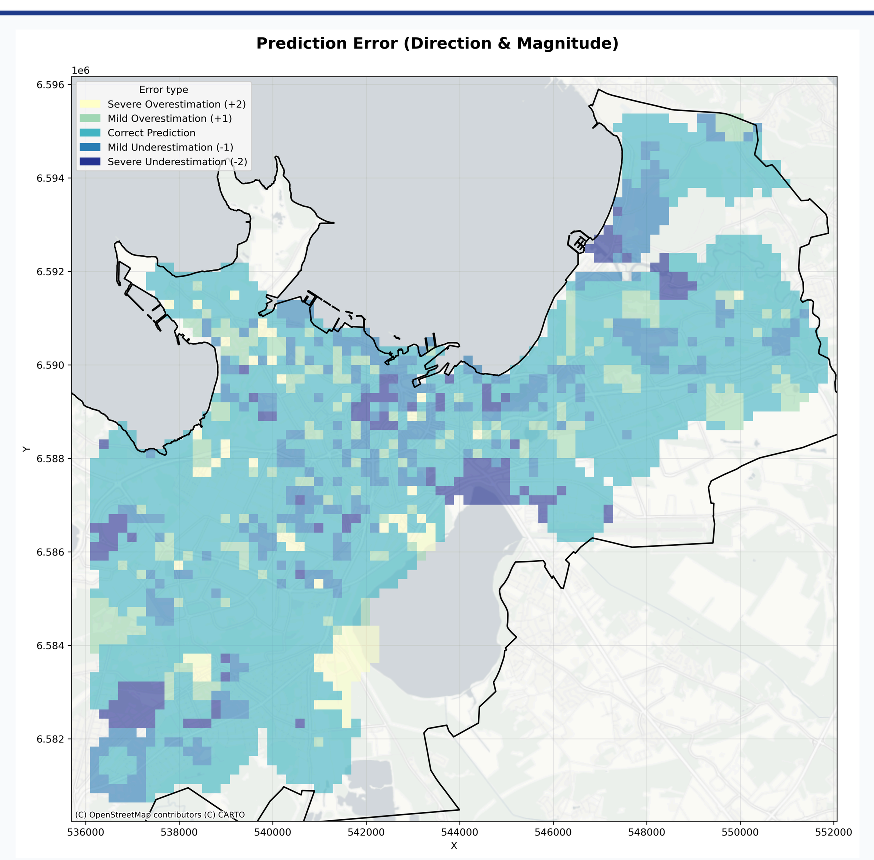
Bivariate Moran's I measures spatial correlation between variables. Key: Price correlates with lower distance to center and higher to primary education. Tallinn is monocentric; sports/leisure cluster with misc services. Weak price segregation.

6. Model Performance

71.4% Accuracy
0.552 Kappa
0.277 R^2

Key features: Target encoding (districts, 2km grids), walkability index, amenity diversity

Challenge: Strong spatial autocorrelation - nearby properties influence each other



Understanding Prediction Error

Color-coded error:

- **Dark Blue:** Underestimation (-2 to -1 class)
- **Blue:** Correct
- **White:** Overestimation (+1 to +2).

Insight: Reveals spatial biases (e.g., overestimation in suburban areas, underestimation near amenities).

7. Conclusions

Spatial autocorrelation dominates: Neighborhood effects strong ($\rho \approx 0.48$ in Tallinn)

Consistent patterns: Proximity to food/drink amenities increases rent

Tartu limitations: With only 448 listings, reliable predictions or assumptions couldn't be made due to insufficient data size

Data Sources: KV.ee • OpenStreetMap • OSMnx • XGBoost • Maa- ja Ruumiamet (Estonian Land and Spatial Development Board)

Repository: Real-Estate-Pricerange-Predictor

Full code and figures available in the project repository

Scan for GitHub