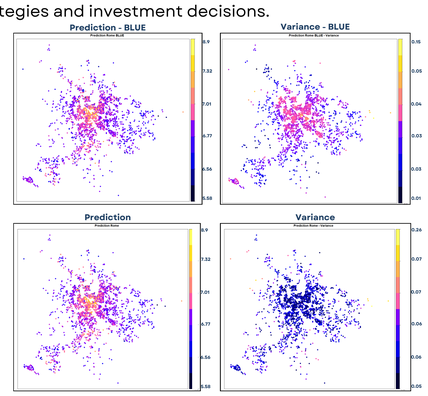
KRIGING

For policy makers

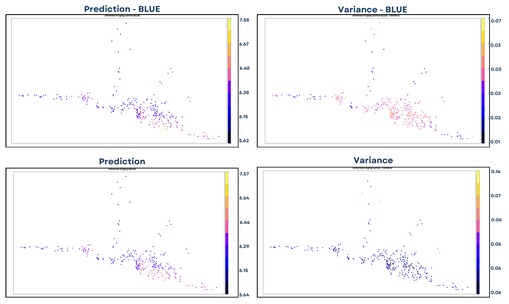
Kriging, a geostatistical interpolation technique, predicts values across a surface, such as house renting prices across a study area. It **leverages known prices at specific points to estimate prices in areas lacking data**. By considering spatial autocorrelation, kriging identifies **spatial patterns and trends in house renting prices**, aiding pricing strategies and investment decisions.

This provides policymakers with empirical evidence and spatial information about renting price dynamics, enabling **data-driven decisions concerning housing policies**. By examining local context and broader spatial trends, policymakers can pinpoint areas with similar characteristics and develop strategies for fair and sustainable housing markets.

Mapping these patterns helps establish fair price benchmarks based on general characteristics, accessibility, and socio-economic factors. Furthermore, monitoring spatial trends in house prices over time allows policymakers to evaluate regulatory measures and implement adjustments to maintain fairness and stability in the housing market, potentially tackling speculative bubbles and promoting equitable access to housing resources.



Kriging - Rome



Kriging -Genoa

Local Moran's Indexes

Bologna	0.5600
Genoa	0.4584
Milan	0.6723
Rome	0.6687
Turin	0.3573

GLS

For landlords

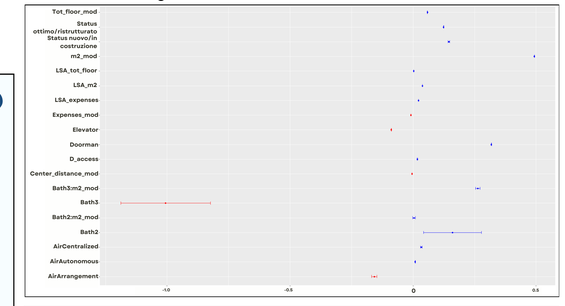
Generalized Least Squares (GLS) estimation is a generalization of the Ordinary Least Squares (OLS) estimation technique, where the assumptions of constant variance and independence of errors are not met. In housing renting prices datasets, it is common to encounter **heteroskedasticity** and **correlation between residuals,** due to different factors such as location, size, neighborhood and other structural characteristics.

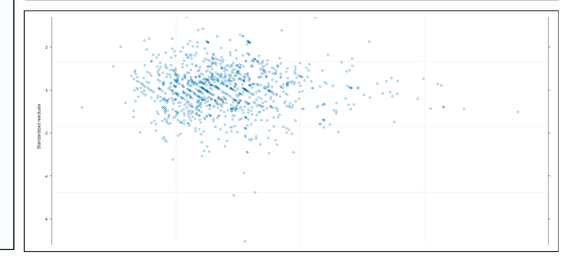
GLS helps in finding the **estimates for model parameters that minimize the overall error**, considering that some points might be more trustworthy than others. To do this, GLS **assigns different weights to each data point based on their amount of variability**. It then adjusts the line to minimize the total weighted error, giving **more importance to the more reliable points** and less to the less reliable ones.

Regarding the dataset of the project, GLS is a useful technique for analyzing relationships between characteristics of houses and their renting prices. Indeed, it can help in understanding how changes in certain features of a house might affect its final renting price.

For sellers or constructors, this knowledge **can be valuable for pricing strategies**, identifying which features contribute the most to the value of a property rent, offering insights on how investments in upgrades or renovations should be done **to yield the highest returns**.

From the gls, it emerges that **relevant variables common to all cities** are m2, total floors in the building, status of the building, energy class, number of bathrooms, expenses. On the other hand, the various cities present **differences in terms of relevant characteristics significantly impacting house prices,** mainly related to geographical and social differences among cities.





GLS - Bologna

WHICH IS A FAIR PRICE FOR YOUR HOUSE?

The aim of this analysis is to examine and interpret dynamics behind the **Italian real estate rental market**, particularly focusing on five main cities: Rome, Milan, Turin, Genoa, and Bologna. The ultimate goal is to develop a **statistical tool** that can **determine a "fair" price for each property**, **benefiting all key stakeholders** in the real estate rental market: constructors/sellers, buyers, and policymakers.

Large cities often experience **disparities in rental prices**, typically disadvantaging buyers. This imbalance can be rectified by **establishing a *fair* price,** which represents the tradeoff between the minimum amount at which the sellers are willing to sell and the maximum at which buyers are willing to buy with respect to the present market scenario. Achieving this **indifference price** is crucial for creating a **mutually beneficial outcome**. Here, policymakers play an important role in regulating the market to prevent powerful actors from exploiting the situation to their advantage.

The data analyzed were collected from *Immobiliare.com* and pertain to **house prices** for both sales and rentals in Italy during the **first semester of 2023**. For our analysis, we specifically focused on **rental data for the aforementioned five cities**. The dataset comprises a total of ***29,362 instances***, encompassing variables such as location, apartment characteristics, energy class, building furniture, etc.

R2	GLS (NO LSA)	GLS (LSA)	KRIGING	RANDOM FOREST (NO LSA)	RANDOM FOREST (LSA)
MILAN	0.7041	0.7143	0.8113	0.8685	0.8864
ROME	0.7435	0.7440	0.8293	0.8714	0.8719
TURIN	0.6381	0.6436	0.1785	0.7061	0.7023
BOLOGNA	0.6359	0.6516	0.4169	0.8029	0.7981
GENOA	0.6141	0.6115	0.6151	0.7015	0.6897

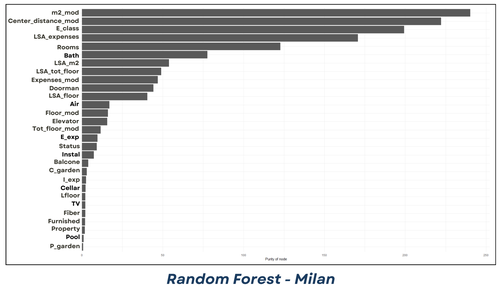
RMSE	GLS (NO LSA)	GLS (LSA)	KRIGING	RANDOM FOREST (NO LSA)	RANDOM FOREST (LSA)
MILAN	0.2681	0.2634	0.2142	0.1797	0.1677
ROME	0.2729	0.2726	0.2226	0.1945	0.1948
TURIN	0.2827	0.2805	0.4125	0.2524	0.2547
BOLOGNA	0.2562	0.2506	0.3100	0.1895	0.1912
GENOA	0.2433	0.2443	0.2417	0.2133	0.2180

RANDOM FOREST

For renters

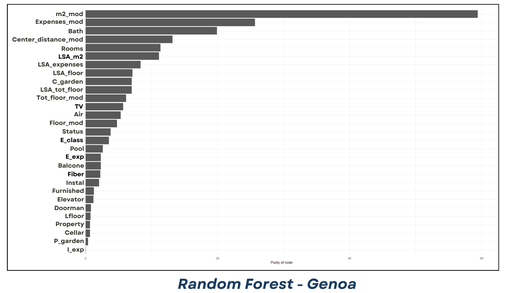
The Random Forest technique, rooted in decision trees, is widely employed for both classification and regression tasks. Essentially, it comprises a "forest" of numerous decision trees, each trained on different subsets of data using bootstrap aggregating (bagging) processes, effectively **reducing variance** and **lowering the risk of overfitting**. Among the models examined so far, it demonstrates the **least variability** in **estimating relevant features that impact house prices** across different cities. Further enhancements are achieved by incorporating local Moran's coefficients.

This model proves to be **particularly valuable for potential renters** who are unfamiliar with local housing markets and pricing dynamics. By **highlighting the impact of different factors on the renting price**, it empowers potential renters to make informed decisions regarding the features they prioritize in a home, enabling them to **find a balance between desired features and budget considerations**. Moreover, the model's **speed** and **greater accuracy** offer many advantages to individuals engaged in the complex task of evaluating numerous properties during their search for a new home.



Random Forest - Milan

TRY NOW!



Random Forest - Genoa

POLITECNICO MILANO 1863

Lorenzo Angeloni 10923468
Luca De Rocchi 10933781
Diego García 10934350
Mattia Longobardo 10616746
Agnese Zoffoli 10936988
MSc Management Engineering, *Applied Statistics, a.a. 2023/24*

References:

- *Gabetti Group. (n.d.). Riqualificazione energetica degli edifici. Un benessere oltre il green. Retrieved February 19, 2024, from https://www.gabettigroup.com/report/https-www-gabettigroup-com-wp-content-uploads-2023-10-polimi-polito-2023-pdf/*
- *Griffith, D. A., Chun, Y., & Li, B. (2019). Spatial regression analysis using eigenvector spatial filtering. San Diego, CA: Academic Press.*
- *Griffith, D. A., & Helbich, M. (2016). Spatially varying coefficient models in real estate: Eigenvector spatial filtering and alternative approaches. Computers, Environment and Urban Systems, 57, 1–11. https://doi.org/10.1016/j.compenvurbsys.2015.12.002*
- *Griffith, D. A., & Chun, Y. (2009). Spatial autoregressive models with eigenvector spatial filters: Spatial econometric methods revisited. Journal of Regional Science, 49(3), 453–479. https://doi.org/10.1111/j.1467-9787.2009.00632.x*