Rapada Partners – CASE STUDY

Assesing the potential of Madrid's Real Estate market.

Brief Summary

The following report delivers a data-driven analysis of Madrid 's property market to highlight key investment opportunities and reduce risk an a constantly changing real estate market.

By separating the sales and rental markets and applying predictive modeling, we provide actionable insights on undervalued assets, district-level trends, and potential returns.

Our findings and conclusions highlight premium districts with strong long-term appreciation, as well as emerging areas offering high gross rental yields. These insights are intended to enable **Rapada Partners** to make informed decisions on where and how to allocate capital for maximum impact.

- ***Objectives***

The goal of this study was to analyze Madrid's real estate market to anticipate trends and detect opportunities that distinguish a profitable investment from a potentian million-dollar loss.

We were provided with a dataset of over 100,000 properties, each with 39 variables containing information on property characteristicas, location, and pricing, as summarized bellow.

Column	Description	Example
ad_url	Link to property listing	https://www.idealista.com/inmueble/102834499/
reference_id	Unique property identifier	102834499
property_type	Type of property	vivienda
ad_title	Ad title	Alquiler de Piso en calle del Amparo, Lavapiés-Emba
ad_description	Ad description	UBK-493404(Disponible 1-11 meses) - Sin posibilida
transaction_type	Sale or rental transaction	alquiler
n_rooms	Number of rooms	1.0
n_baths	Number of bathrooms	1.0
area	Surface area (m^2)	58.0
floor	Floor number	1.0
has_elevator	Elevator availability	False
has_terrace	Terrace availability	False
is_exterior	Exterior property?	False
has_swimming_pool	Swimming pool availability	False
has_parking	Parking availability	False
has_garden	Garden availability	False
energy_certificate_consumption	Energy consumption rating	
energy_certificate_emissions	Energy emissions rating	
energy_certificate_state	Energy certificate state	UNKNOWN
has_energy_certificate	Energy certificate available?	True
has_floorplan	Has floorplan?	False
has_virtual_tour	Has virtual tour?	False
n_videos	Number of videos	0
property_state	Property state	Buen estado
latitude	Latitude	40.4084991
longitude	Longitude	-3.7007005
approximate_location	Approximate location available?	True
address	Property address	Plaza de Lavapiés
province	Province	Madrid
country	Country	España
country_code	Country code	es
state	Region	Comunidad de Madrid
city	City	Madrid
postcode	Postal code	28012.0
district	District	Centro
quarter	Quarter	Embajadores
price_down_from	Original price (if discounted)	nan
price	Current price	2067.0
viewed_at	Date when ad was last viewed	2024-08-01

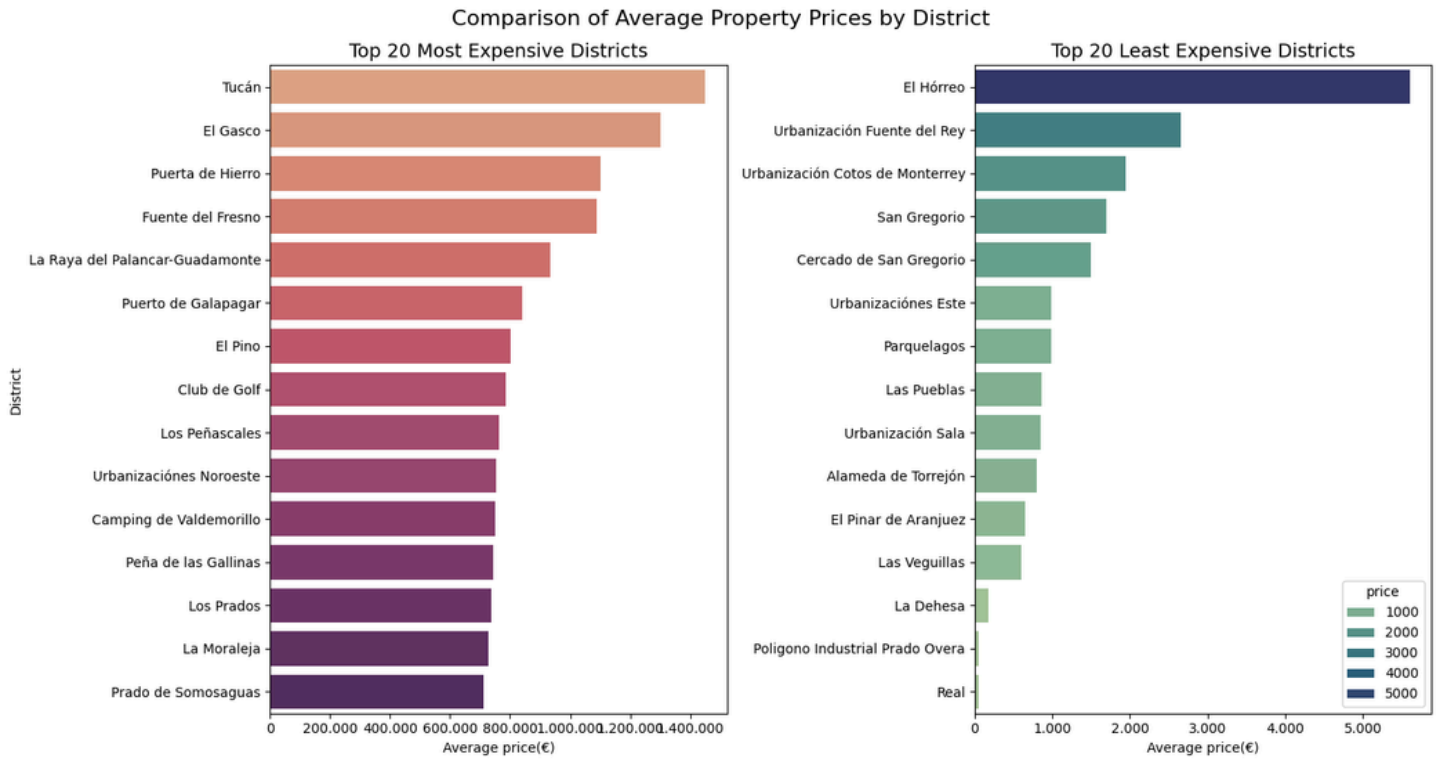
Using advanced analytical techniques and widely recognized AI models, we identified undervalued properties and high-yiedl districts to hep strategic investment decisions.

- Data and Methodology

We first prepared the data by cleaning missing values and removing outliers in key variables such as properties with anomalous areas or prices.

To provide context for the analysis, we first explored the distribution of property prices and characteristics across Madrid's districts.

The following visualizations illustrate the significant variability in average property prices between central premium areas and peripheral districts, highlighting the importance of location as a key driver in Madrid's real estate market. Additionally, we examined price distributions to identify and remove extreme outliers that could bias the predictive models.



We also engineered new features by performing calculations on existing dataset fields:

- Price per m²: capturating local price density
- “luxury” indicator: selecting properties > €1M and > 150 m².
- Gross rental yield: estimated as annual rent divided by pruchase price

- Market Segmentation

To capture the different dynamics of sales and rentals.

- Sales model predicts property prices and highlights undervalued assets for acquisition
- Rental model predicts monthly rents and enables estimation of gross rental yield

- Predictive Modeling

To predict property prices in both the sales and rental markets, we applied Random Forest Regressors. This algorithm is widely used in real estate analytics for its ability to capture complex patterns and variable interactions , such as the combined influence of property size, location, and amenities on pricing.

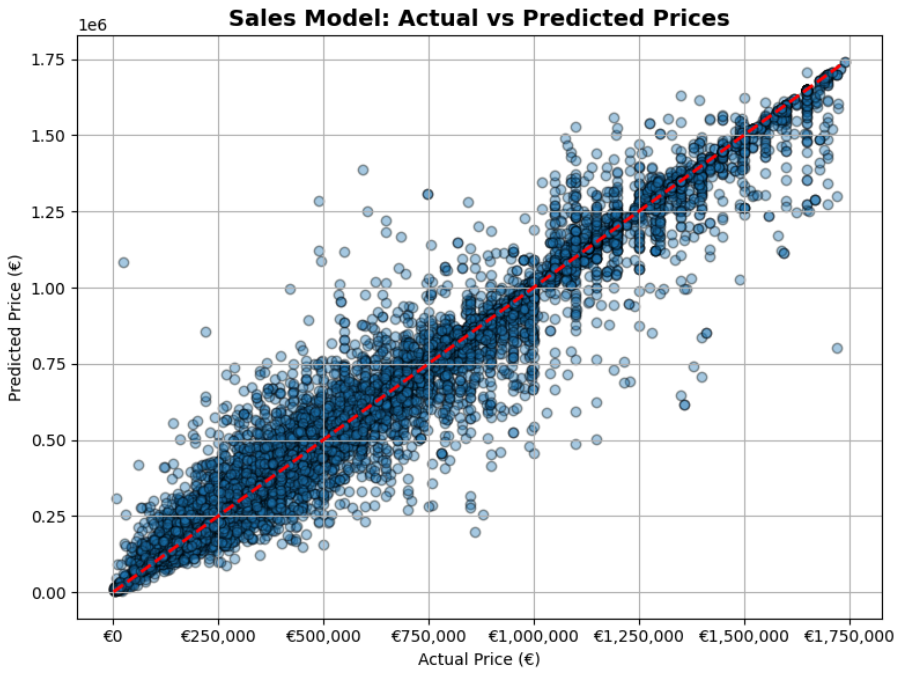
In simple terms, Random Forest builds multiple decision trees and combines their outputs to produce a more accurate and robust prediction. This approach helps minimize errors and ensures the model performs well even with diverse property types and price ranges.

To evaluate the performance of our predictive models, we used two widely recognized metrics: **Mean Absolute Error (MAE)** and the **R² Score**.

- **Mean Absolute Error (MAE):** This metric measures the average difference between the predicted prices and the actual prices in the dataset. For property sales, the model's predictions deviate by an average of **€30,661**, and for rentals, by **€224**. While these numbers may appear large at first, they are relatively small compared to the typical property prices in Madrid. This suggests that the model provides predictions that are precise enough to guide investment decisions effectively.
- **R^2 Score (Coefficient of Determination):** This measures how well the model captures and explains the variability in property prices. A score of **1.0** would indicate perfect predictions. Our models achieved an R^2 of **0.97** for sales and **0.88** for rentals. These values show that the models successfully capture most of the underlying patterns in the data, making them highly reliable tools for identifying trends and assessing potential investments.

By combining these metrics, we can say that our models provide a solid foundation for evaluating property values and estimating return.

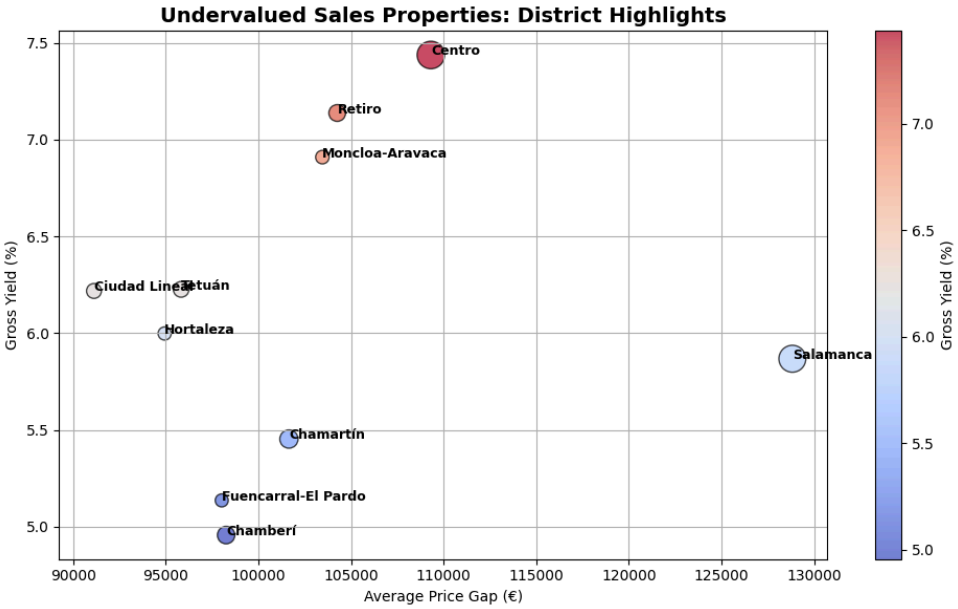
The scatter plot illustrates the accuracy of the sales price predictions: most points cluster closely around the diagonal red line, indicating a strong alignment between actual and predicted prices.

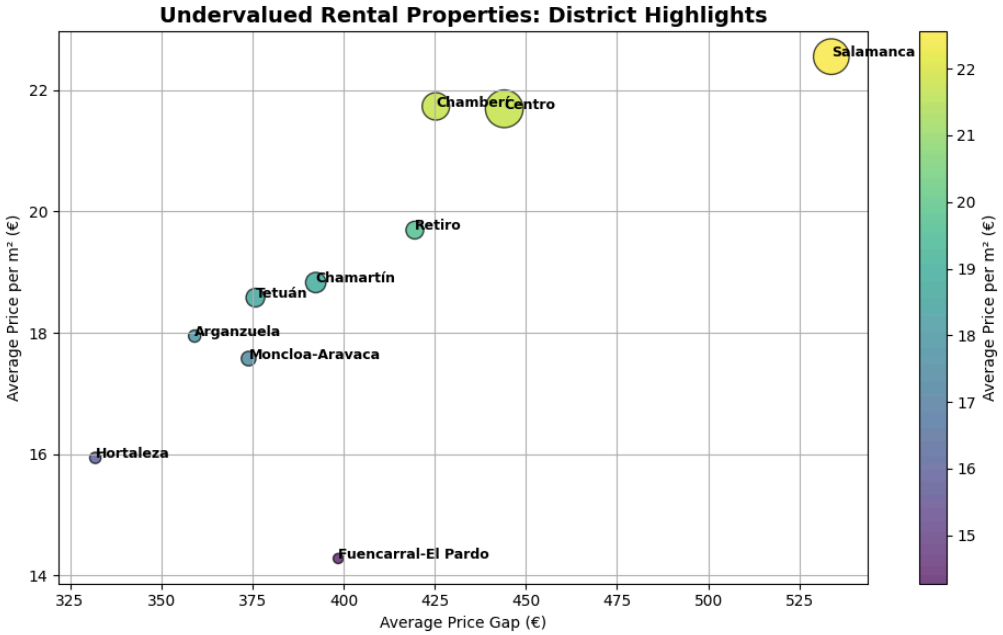


- Key Findings
  - **Sales Prices:** are primarily driven by luxury status, property area, and district. Salamanca, Chamartín, and Centro are the most influential locations.
  - **Rental Prices:** are mainly determined by property area, number of bathrooms, and energy certification.

The analysis of undervalued properties in both the sales and rental markets reveals clear patterns across Madrid's districts.

The following visualizations highlight the districts with the highest number of undervalued properties, average price gaps, and potential returns.


Undervalued Sales Properties: District Highlights
Centro
Retiro
Moncloa-Aravaca
Ciudad Lineal
Tetuán
Hortaleza
Salamanca
Chamartín
Fuencarral-El Pardo
Chamberí
Gross Yield (%)
Average Price Gap (€)
7.5
7.0
6.5
6.0
5.5
5.0
90000
95000
100000
105000
110000
115000
120000
125000
130000


Undervalued Rental Properties: District Highlights
Salamanca
Chamberí
Centro
Retiro
Chamartín
Tetuán
Arganzuela
Moncloa-Aravaca
Hortaleza
Fuencarral-El Pardo
Average Price per m² (€)
Average Price Gap (€)
22
20
18
16
14
325
350
375
400
425
450
475
500
525