

Model Agnostic Methods

Prepared for: UPV Renting Bike .SA

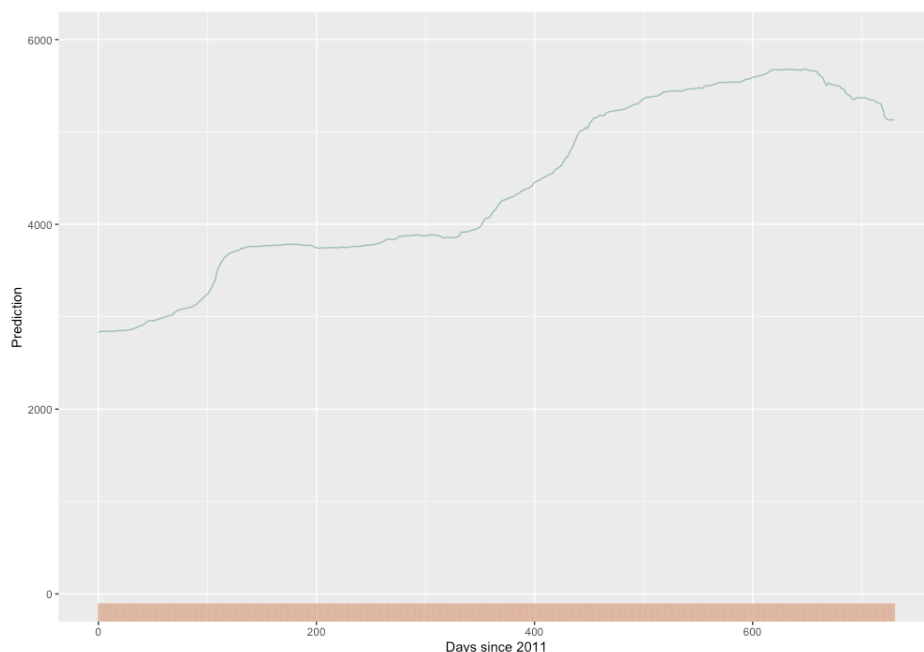
Prepared by: Adrián Tallec de León, Jezabel Esbrí Rodríguez, Manuel Rocamora Valenti.

Introduction

In this report, we apply model agnostic methods to understand the influence of various factors on predicted outcomes. Specifically, we utilize Partial Dependence Plots (PDPs) and Bidimensional Partial Dependency Plots to visualize the relationships learned by Random Forest models. This analysis is crucial for interpreting the effects of environmental, temporal, and property features on bike rental counts and house prices. Our objective is to provide actionable insights for business operations, urban planning, and real estate valuation.

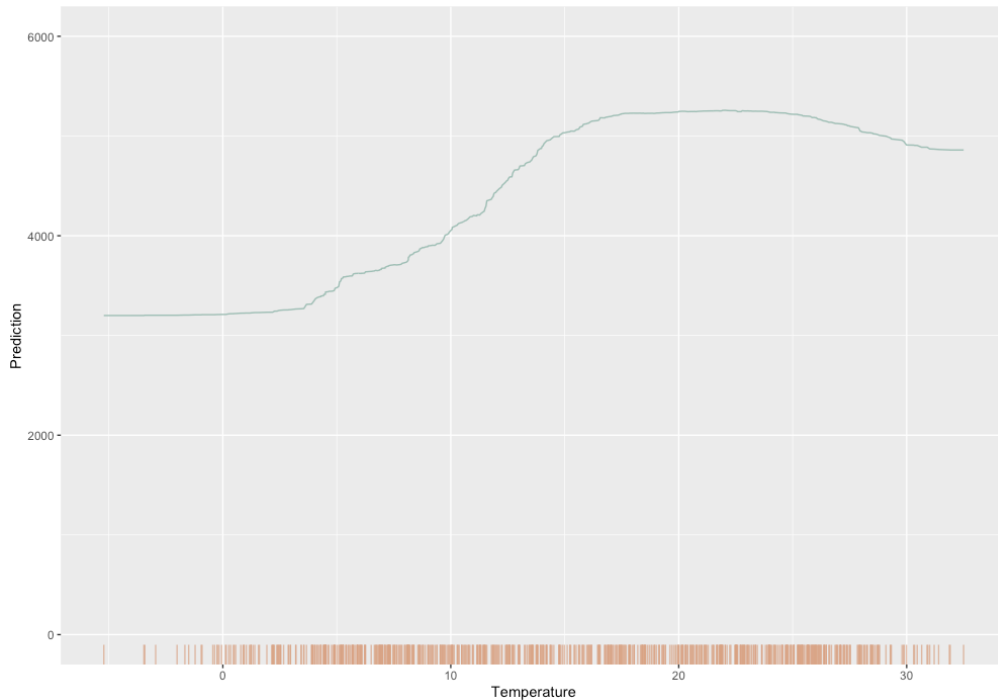
Partial Dependence Plot

To analyze the influence of key environmental and temporal variables—days since 2011, temperature, humidity, and wind speed—on predicted bike rental counts, we utilized a Random Forest model visualized through Partial Dependence Plots (PDP). Each of these factors was systematically examined to understand their individual impacts on the frequency of bike rentals, as follows.



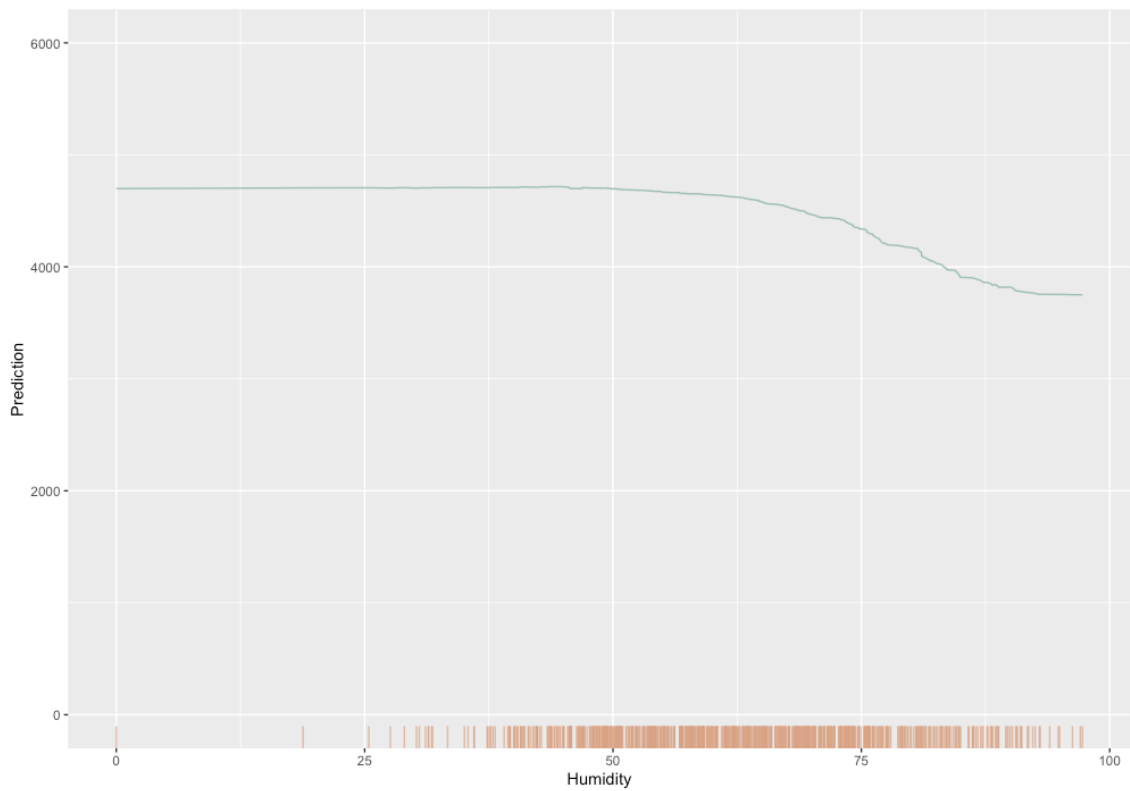
Days Since 2011:

The PDP for days since 2011 shows a general upward trend in bike rentals over time, peaking around 600 days before experiencing a slight decline. This indicates an increase in the popularity or availability of bike rentals over the years, with a potential saturation or operational challenge impacting the rentals after the peak period.



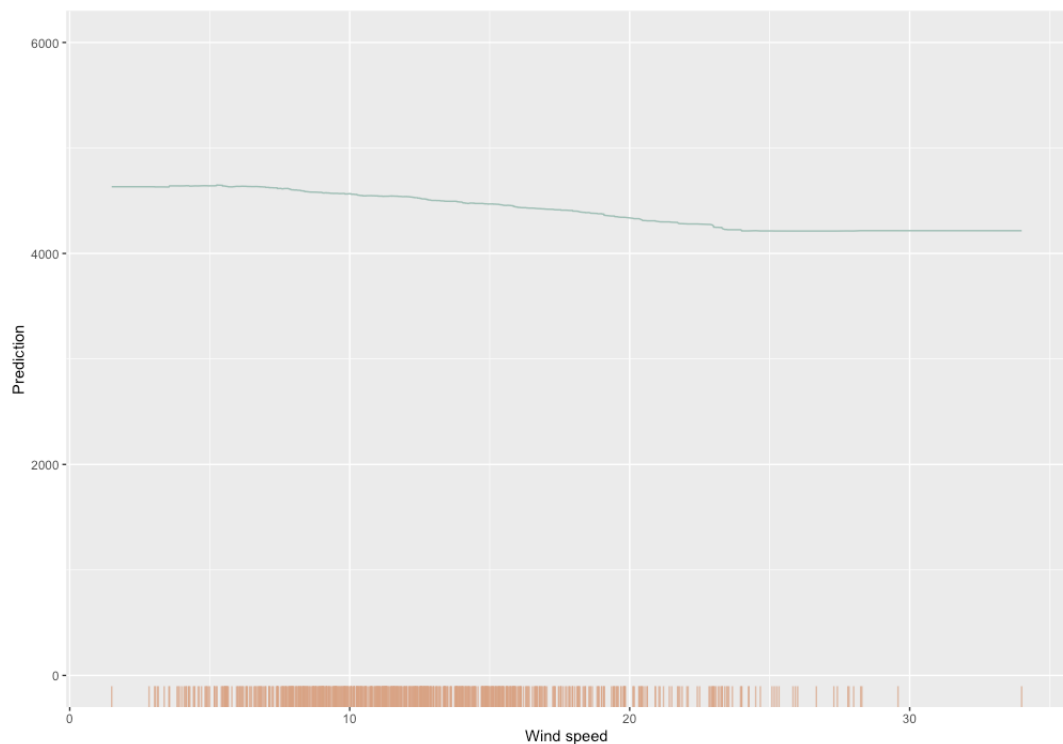
Temperature:

The temperature plot reveals a strong positive relationship between temperature and bike rentals up to about 20 degrees Celsius, after which the trend plateaus. This suggests that warmer weather up to a comfortable point encourages more bike usage, likely due to more favorable biking conditions. However, temperatures above this point do not significantly increase rental counts, possibly due to already sufficient warmth for biking or other limiting factors at very high temperatures.



Humidity:

The influence of humidity on bike rentals shows a consistent decline as humidity increases. Higher humidity levels, which often correlate with discomfort and poorer biking conditions, appear to deter bike rentals. This relationship is quite linear, indicating that even small increases in humidity can reduce the propensity for renting bikes.

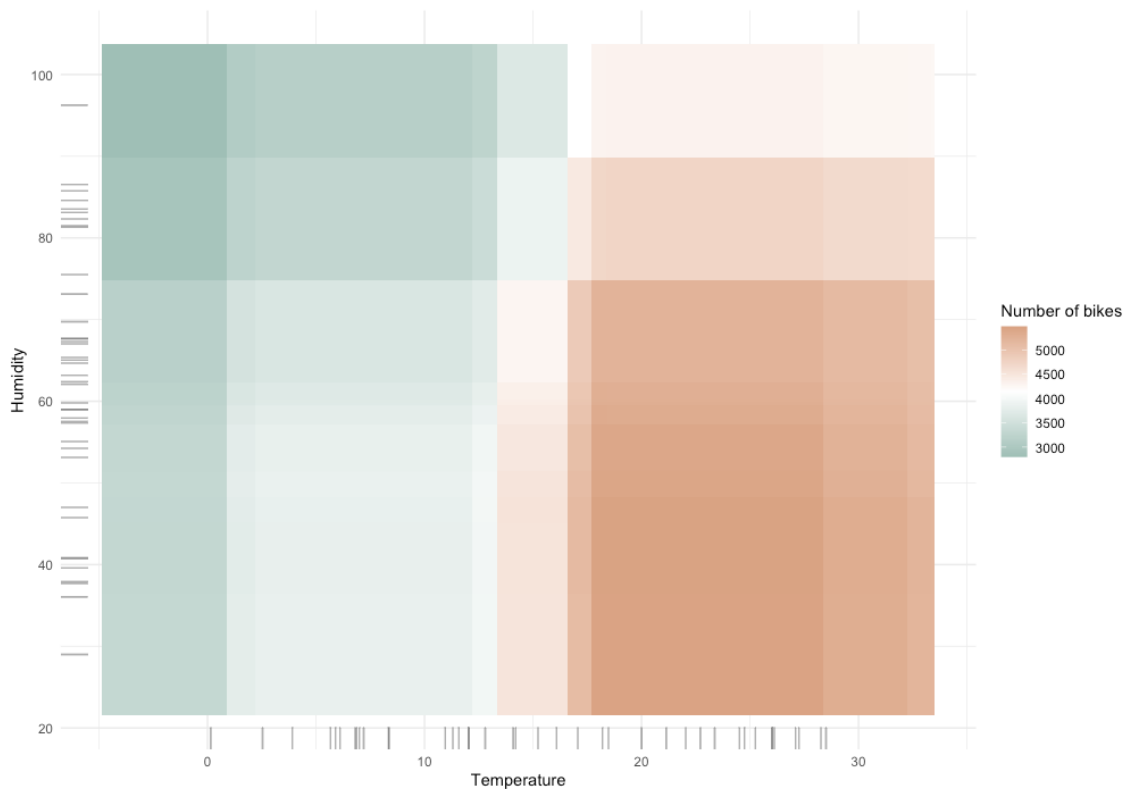


Wind Speed:

The influence of humidity on bike rentals shows a consistent decline as humidity increases. Higher humidity levels, which often correlate with discomfort and poorer biking conditions, appear to deter bike rentals. This relationship is quite linear, indicating that even small increases in humidity can reduce the propensity for renting bikes.

Bidimensional Partial Dependency Plot

In this section of the report, we analyze the interplay between temperature and humidity and its effect on bike rental patterns. Utilizing a 2D Partial Dependency Plot with a Density Overlay, we can visually decipher and quantify how these two climatic factors influence bike rental behavior. This analysis not only helps in understanding the dynamics of bike rentals but also aids in making informed decisions for business operations and urban planning.



Effects of Temperature and Humidity:

The plot reveals that the predicted number of rented bicycles varies significantly with changes in temperature and humidity.

There is a gradient of predicted bike rentals that increases as the temperature rises, particularly notable in the mid-range of humidity values (approximately 50% to 75%). This suggests that warmer temperatures generally increase the likelihood of higher bike rentals, especially when humidity is moderate.

Impact of Humidity:

Higher humidity levels (above 75%), combined with any temperature, tend to correlate with a decrease in bike rentals. This is observed in the transition to cooler colors in areas of higher humidity on the plot.

The lower humidity range (below 50%) shows a mixed impact but tends to have higher rentals at higher temperatures, indicated by warmer colors in the plot at these points.

Optimal Conditions for Bike Rentals:

The plot suggests that the optimal conditions for bike rentals, in terms of maximum predicted numbers, occur at higher temperatures (around 20° C to 30° C) with moderate humidity (around 50% to 75%).

The least favorable conditions for bike rentals appear to be when both temperature and humidity are very high, or when humidity is very high, regardless of temperature.

Data Distribution:

The density contour lines overlaid on the plot show where most of the observed data points cluster. High-density areas are marked with contour lines closer together.

There is a noticeable concentration of data points at mid to high temperatures and mid-level humidity, suggesting that these conditions are common in the dataset. This might also reflect typical weather conditions during the data collection period.

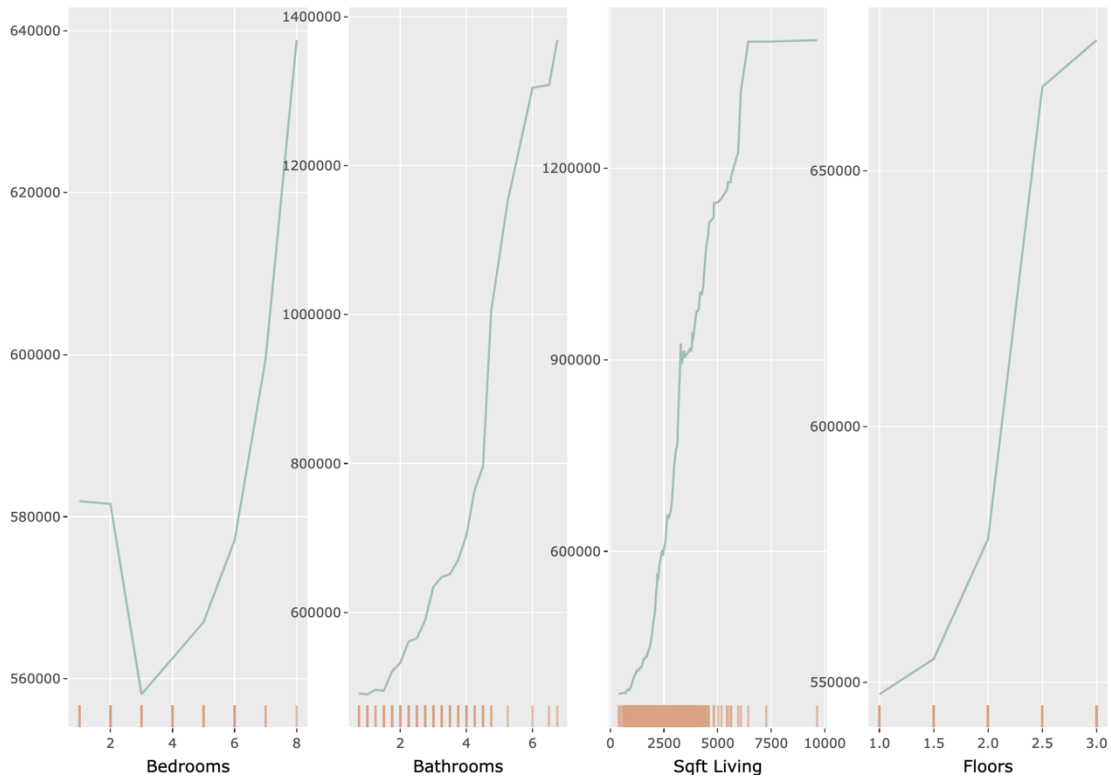
Density and Predictions:

Regions with the highest density of data points do not always align with the highest predicted rentals, indicating that while certain weather conditions are common (warm and moderately humid), they do not always maximize bike rentals.

Conversely, some areas with fewer data points (less dense) correspond to higher or lower predictions of bike rentals. This could suggest less frequent but impactful weather conditions on rental behavior.

PDP Explain Price Of House

Understanding the factors that influence house prices is essential in the real estate market. This analysis uses a Random Forest model to predict house prices based on key features: bedrooms, bathrooms, square footage of living space (sqft_living), and the number of floors. By generating Partial Dependency Plots (PDPs), we visualize the relationships between these features and house prices, providing valuable insights for real estate professionals, buyers, and sellers.



Bedrooms

The plot for bedrooms shows a complex relationship:

- Prices decrease as bedrooms increase from 1 to 2.
- There's a sharp increase in price as the number of bedrooms goes from 2 to around 4, then a steep decline with 5 bedrooms, followed by another peak at 6 bedrooms.
- The trend suggests that both very small (1–2) and larger (6+) bedroom counts are associated with higher prices, possibly due to the market demand for either compact homes or larger, more luxurious properties.

Bathrooms

The relationship between the number of bathrooms and house prices shows a more consistent trend:

- Prices increase almost linearly with the number of bathrooms, suggesting that additional bathrooms significantly enhance property value.
- This could reflect the utility and luxury of having multiple bathrooms, which is often a selling point for larger or more upscale homes.

Square Feet of Living Area (sqft_living)

This plot indicates a strong positive relationship between the living area in square feet and the house price:

- There is a sharp, steady increase in price as the living area increases, reflecting the premium placed on larger homes.
- This trend is consistent with general real estate principles where larger living spaces command higher prices due to increased utility and appeal.

Floors

The floors plot reveals that:

- House prices increase with the number of floors, with a significant rise observed from 1 to 2 floors and a slighter slope from 2 to 3 floors.
- This could indicate that multi-story homes are valued higher, possibly due to better views, separation of living spaces, and the perception of luxury.

Conclusion

This report demonstrates the application of model agnostic methods to understand the influence of various features on predicted outcomes for bike rentals and house prices. The Partial Dependence Plots (PDPs) and Bidimensional Partial Dependency Plots provide valuable insights into how key environmental, temporal, and property features impact these predictions.

Understanding these relationships is crucial for optimizing business operations, urban planning, and real estate valuation, allowing stakeholders to make informed decisions based on the data-driven analysis presented.