# Model Agnostic Methods

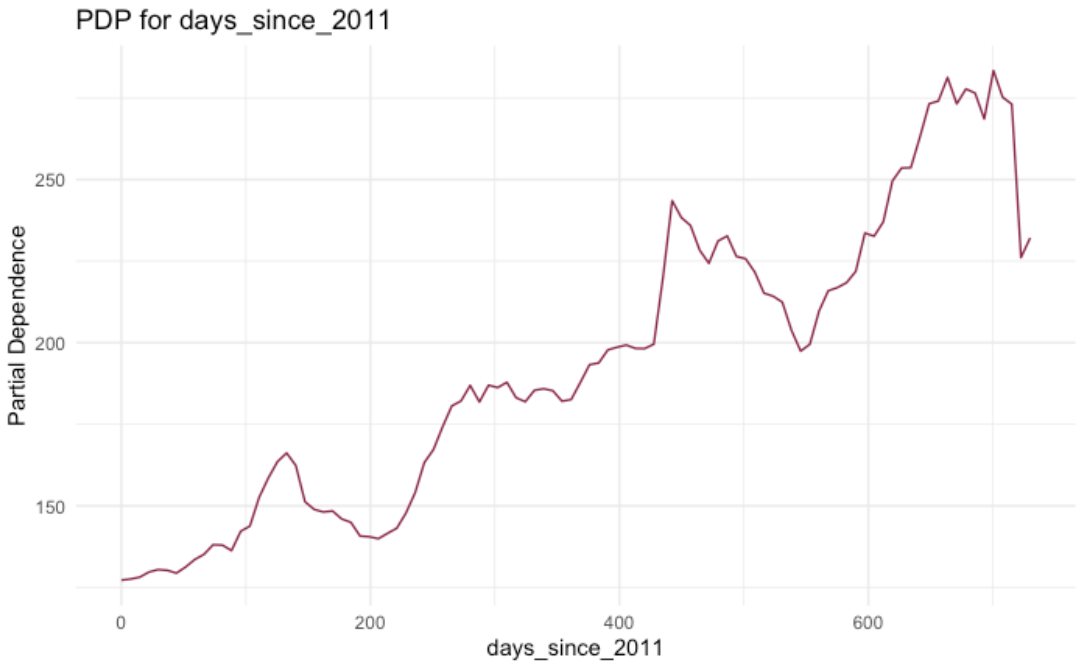
**Prepared for**: UPV Renting Bike .SA
**Prepared by**: Adrián Tallec de León, Jezabel Esbrí Rodríguez, Manuel Rocamora Valenti.

## 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.

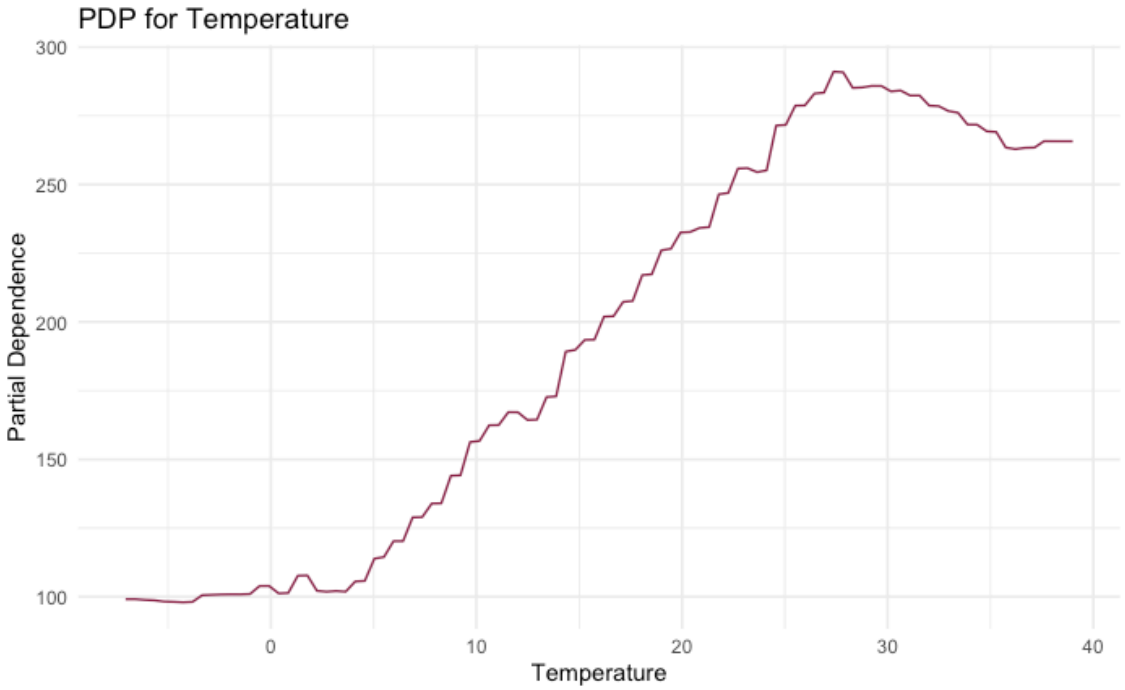
## Partial Dependence Plot

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.



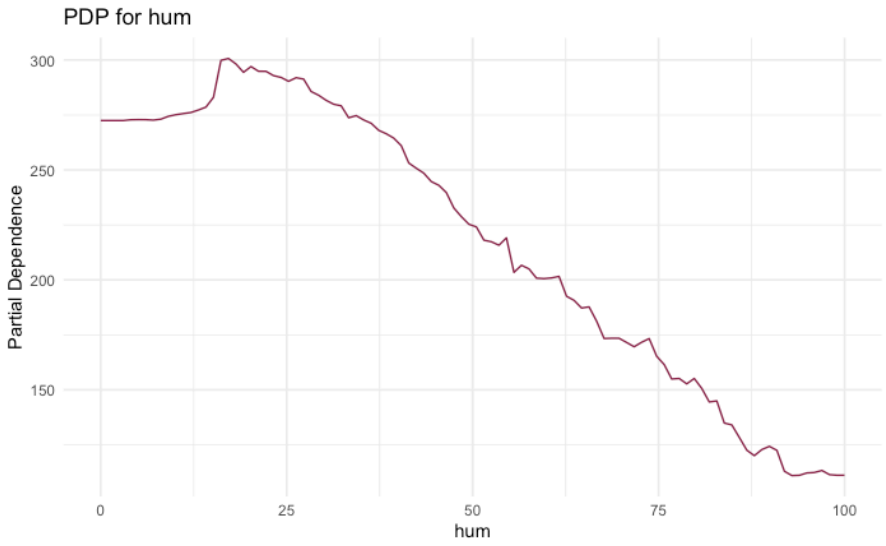
Days Since 2011:

The PDP shows a fluctuating pattern, which suggests that the number of days since 2011 has a non-linear relationship with bike rentals. Peaks and troughs indicate that certain periods are more associated with higher rentals, which could be linked to seasonal effects, specific events, or changes in biking infrastructure over time.




Temperature:

The plot for temperature shows an increasing trend in bike rentals as temperature rises, up to a certain point, after which it starts to plateau and slightly decline. This indicates that there is an optimal temperature range for bike rentals, which makes sense as extremely high temperatures might deter outdoor activity.



### Humidity:

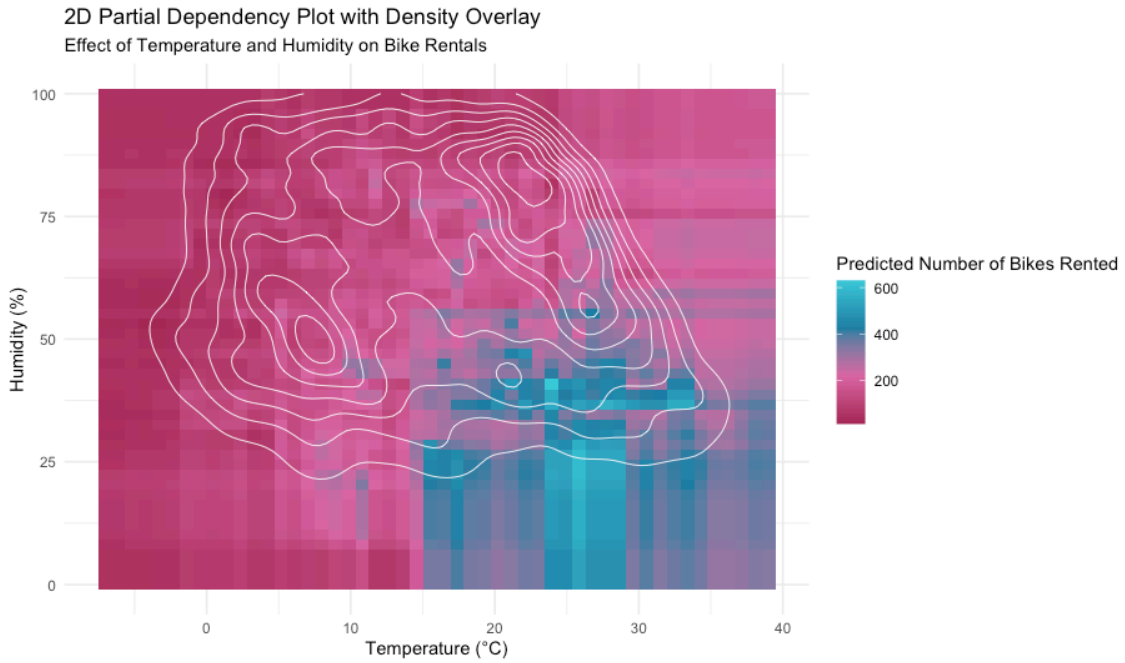
The influence of humidity on bike rentals is relatively clear, with higher humidity levels leading to a consistent decrease in bike rentals. This trend suggests that less comfortable conditions (higher humidity) dissuade people from renting bikes.


### Wind Speed:

The plot for wind speed shows that as the wind speed increases, bike rentals slightly decline after remaining stable up to a certain point. This pattern could indicate that moderate wind speeds do not significantly impact bike rentals, but higher winds might negatively affect the decision to rent bikes due to increased pedaling effort or safety concerns.

## 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

- General Trend: The predicted house price generally increases with the number of bedrooms. This indicates that larger homes with more bedrooms are typically valued higher.

- Step Changes: Notable step increases in price are observed around 3 and 6 bedrooms. This suggests significant price jumps when crossing these thresholds, reflecting a higher valuation for homes with more rooms, which can accommodate larger families or offer more utility.

- Saturation Point: Beyond 6 bedrooms, the increase in predicted price levels off. This saturation point suggests that additional bedrooms beyond this number do not significantly contribute to further increases in house value, possibly because the market demand for homes with such a high number of bedrooms is relatively niche.

### Bathrooms

- Smooth Increase: The predicted house price shows a smooth and continuous increase with the number of bathrooms. More bathrooms add convenience and comfort, contributing to higher property values.

- Notable Jumps: Significant price increases are seen when moving from 2 to 3 bathrooms and again from 4 to 5 bathrooms. These jumps may correspond to market perceptions of luxury and improved living standards associated with having multiple bathrooms.

- High Values: Houses with more than 5 bathrooms are predicted to have significantly higher prices, indicating that these are likely luxury properties with high-end amenities.

### Sqft_living (Square Footage of Living Space)

- Linear Increase: The predicted house price increases almost linearly with the square footage of living space. Larger living spaces offer more comfort and utility, thus driving higher prices.

- Rapid Rise: A noticeable rapid increase in predicted price is observed as the living space exceeds 3000 sqft, with a steeper rise continuing up to around 8000 sqft. This sharp increase highlights the premium placed on larger homes.

- High-End Properties: Homes with over 8000 sqft of living space see a significant increase in predicted price, underscoring their categorization as high-value properties typically associated with luxury and extensive amenities.

### Floors

- Step Changes: The predicted house price shows step changes with the number of floors. A significant jump in price is observed when increasing from 1 to 2 floors, reflecting the higher value placed on multi-story homes.

- Stabilization: The price stabilizes between 2 to 3 floors, suggesting that the major price increases are captured within the first two floors. Beyond this, additional floors do not significantly enhance the property's value.

- **Incremental Increase**: There is a slight increase when moving from 2.5 to 3 floors, but this is not as pronounced as the initial jump from 1 to 2 floors, indicating diminishing returns from additional floors beyond a certain point.

# 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.