Model-Agnostic Interpretability with Partial Dependence Plots: Applications in Bike Rentals and House Price Prediction

Laiqian Ji
Ernesto Martínez Gómez
Álvaro Prado Expósito

Friday 16$^{\text{th}}$ May, 2025

The code developed for this task is contained in a GitHub repository, which can be explored in this link.

1 Introduction

As machine-learning (ML) systems permeate high-stakes domains—healthcare, finance, energy, and public policy—their predictions can no longer be accepted at face value. Stakeholders must be able to trace a model's reasoning, verify that it aligns with domain knowledge, and guard against spurious correlations or discriminatory behaviour. Explainable Artificial Intelligence (XAI) offers a methodological and algorithmic toolbox for translating the opaque logic of complex "black-box" models into forms that are intelligible to human experts and lay users alike.

Among the repertoire of model-agnostic XAI techniques, **Partial Dependence Plots** (PDPs) have proved especially valuable. A PDP depicts the marginal effect of one (or a pair of) explanatory variables on the model's predicted response, averaging out the influence of all other features. By visualising these ceteris-paribus relationships, analysts can infer whether the learned function is monotonic, non-linear, or exhibits interaction effects, thereby gaining practical insight into how changes in specific inputs propagate to the output.

This study applies PDPs to two representative regression problems drawn from open, real-world datasets:

1. **Bike-sharing demand.** Hourly rental counts are predicted from meteorological conditions (temperature, humidity, wind speed) and temporal indicators (season, weekday, hour).

2. **Residential property valuation.** Sale prices of detached houses are modelled as a function of structural attributes (living area, number of rooms, age) and temporal market signals.

For each task we fit an ensemble of 500-tree **Random Forest** (RF) regressors, chosen for their competitive predictive performance, robustness to multicollinearity, and natural compatibility with PDP analysis. Both one-dimensional and two-dimensional PDPs are constructed for the most influential predictors, as determined by permutation importance. The resulting visualisations reveal, for example, threshold effects in temperature on bike demand and non-linear price premiums associated with living area in the housing market.

2 One–Dimensional Partial Dependence Analysis: Bike Rentals

2.1 Dataset and Predictive Objective

We analyse the publicly available *Capital Bikeshare* dataset released by Fanaee - T and Gama (2013), which records hourly rental transactions collected in Washington, DC. Each observation comprises the rental count (`cnt`) together with concurrent meteorological (`tmp`, `hum`, `windspeed`), calendar variables and other variables. The present experiment seeks to quantify, through Partial Dependence Plots (PDPs), the marginal influence exerted by four candidate predictors:

- `d_2011` – integer number of days elapsed since 1 January 2011;
- `temp` – normalised air temperature (in ^{o}C);
- `hum` – relative humidity (fraction);
- `windspeed` – normalised wind speed.

2.2 Model Architecture and Training Protocol

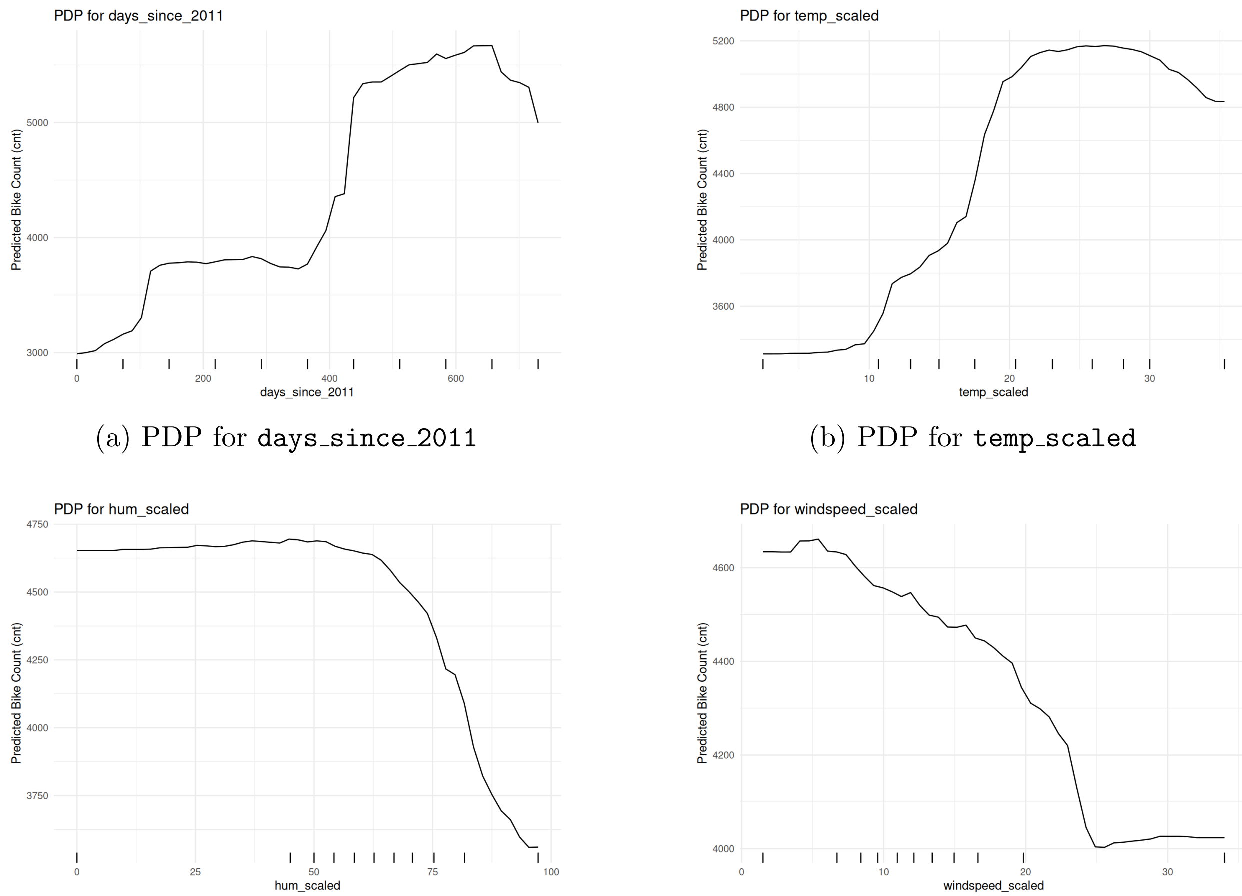
The raw daily rental table (`day.csv`, 731 records) is first enriched with domain-specific, human-readable covariates:

- **Seasonal dummies**. Three binary flags—`spring`, `summer`, `fall`—are derived from the integer `season` code so that winter is the implicit reference category.
- **Weather condition dummies.** `MISTY` identifies "misty" hours (`weathersit`=2); `RAIN` aggregates rainy, snowy or stormy hours (`weathersit`$\in 3, 4$).
- **Physical scaling.** The normalised meteorological inputs supplied with the data are converted back to physical units via simple linear rescalings $temp\_scaled = 41 \times temp$, $hum\_scaled = 100 \times hum$, $windspeed\_scaled = 67 \times windspeed$.
- **Temporal trend.** `days_since_2011` counts the exact number of calendar days elapsed since 1-Jan-2011, enabling the forest to capture long-term demand drift.

All engineered fields plus the original business-cycle indicators (`workingday`, `holiday`) are bundled into a modelling frame named `model_data_day`. After verification that no missing values are present, the regressor is set up.

In this case, a 500-tree **Random Forest** regressor is fitted with the `ranger` package. The forest ingests all 731 observations and returns a regressor with $R^2 = 0.88$ for the training data.

2.3 Interpretation of Results



(a) PDP for `days_since_2011`

(b) PDP for `temp_scaled`

(c) PDP for `hum_scaled`

(d) PDP for `windspeed_scaled`

Figure 1: Exercise 1 plots

- **Days since 1-Jan-2011.** The partial-dependence curve slopes upward almost monotonically: predicted demand grows from just above 3 000 daily rides at the start of the study window to close to 5 000 by late-2012. This secular trend, isolated from weather influences, reflects the rapid adoption of the bike-sharing scheme over its first two years. We can also identify two moments of critical growth and a small decrease at the very end of the interval.

- **Temperature (ºC).** A strong, nearly linear positive effect is visible between about 10 ˚C and 25 ˚C—the comfort band for casual cycling—after which the curve starts to flatten. Warmer weather therefore boosts ridership up to a point, but the marginal gain tapers once summer heat exceeds roughly 30 ˚C.
- **Relative Humidity (%).** The humidity PDP traces a shallow inverted-U: rental counts are highest at low to mid range humidity (< 60 %) and fall away in very muggy conditions (> 80 %). Extreme humidity thus discourages usage, whereas moderate levels have little penalty.
- **Wind Speed (km h^{-1}).** Predicted demand declines steadily as wind intensifies. Moving from calm days to winds near 30 km h^{-1} trims a few hundred trips from the expected total, confirming wind as a deterrent for potential riders.

3 Bidimensional Partial Dependence Plot – Bike Rentals

3.1 Model Architecture and Training Protocol

The second experiment switches from `day.csv` to the far denser `hour.csv` table (with 17 379 observations). Pre-processing mirrors the daily pipeline but is executed at the hourly grain:

- **Physical units.** Normalised meteorological fields are rescaled to their physical ranges (`temp_scaled` = 41 × `temp`, `hum_scaled` = 100 × `hum`, `windspeed_scaled` = 67 × `windspeed`).
- **Categoricals.** `season`, `weathersit`, `holiday`, `workingday`, `weekday` and `yr` are cast to factors so that the forest can isolate non-linear calendar effects.

3.1.1 Sub-sampling strategy

To keep the grid-based PDP computation tractable, a contiguous block of 1.000 rows is drawn at random (start index 7452) from the chronologically ordered data. Sampling a block rather than independent rows preserves local temporal correlations that would otherwise be broken.

3.1.2 Random-Forest fit.

A 500-tree regression forest is trained on the 1 000-row slice with default hyper-parameters. The model reports an $R^2 = 0.778$, confirming adequate fidelity despite the reduced sample.

3.2 Interpretation of Results

Using the `pdp` package we compute a joint partial-dependence surface for `temp_scaled` and `hum_scaled`. The resulting grid is rendered with `geom_tile()`, and marginal rugs of the observed feature values are added for reference. Tile width/height are estimated from the average spacing of the grid to prevent visual “holes”.

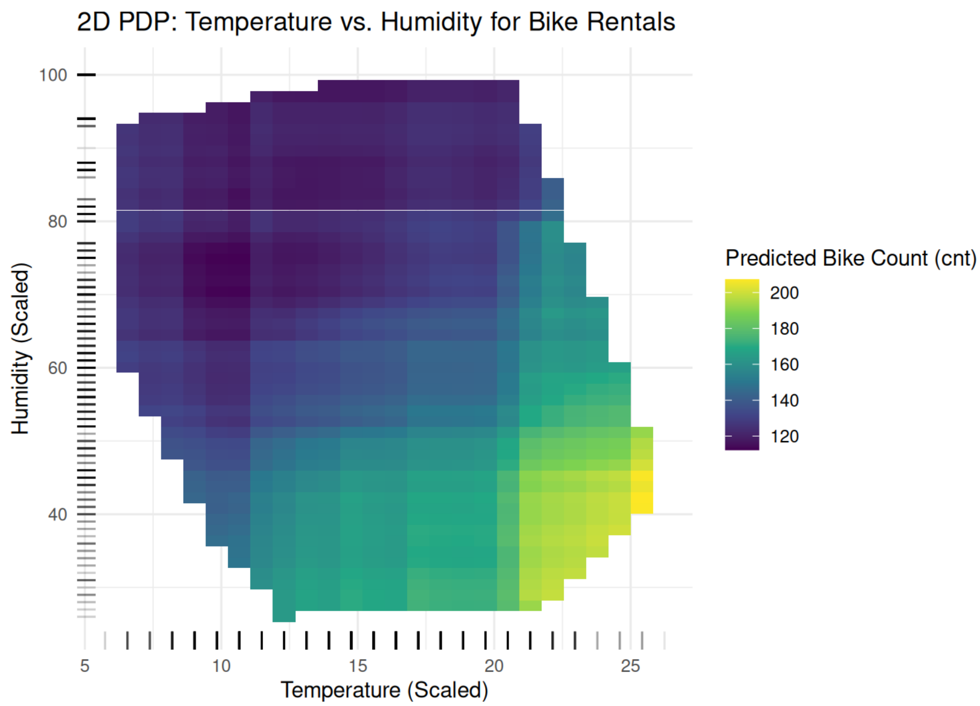


Figure 2: 2D-PDP for `Temperature` and `Humidity`

- **Primary temperature effect.** Across all humidity levels, warmer air drives higher hourly rentals: predicted counts climb steeply from ≈ 120 rides at 5 ºC to over 180 rides near 25 ºC. Temperature is therefore the dominant driver in the interaction.
- **Moderate-humidity effect.** The brightest band on the heat-map—indicating peak demand—occurs where relative humidity sits between roughly 40% and 60% and temperature exceeds 15 ºC. Within this area, marginal temperature gains translate into higher usage.
- **Penalty from high humidity.** At humidities above 80% the surface slopes downward, decreasing predicted rides even when temperatures remain favourable. Sticky, oppressive conditions thus decrease demand independently of warmth.
- **Cold-and-wet trough.** The lowest expected counts appear in the upper-left quadrant ($\leq 10C$, $\geq 70\%RH$), confirming that combined cold and high moisture is particularly discouraging for potential riders.

4 One-Dimensional Partial Dependence Analysis: House Princing

4.1 Dataset and Objective

The third experiment utilises the `kc_house_data.csv` file (21 columns, 21613 sales) that records single–family home transactions in King County, Washington. Our objective is to predict the `price` (continuous target) from a compact subset of structural attributes: `bedrooms`, `bathrooms`, `sqft_living`, `sqft_lot`, `floors` and `yr_built`.

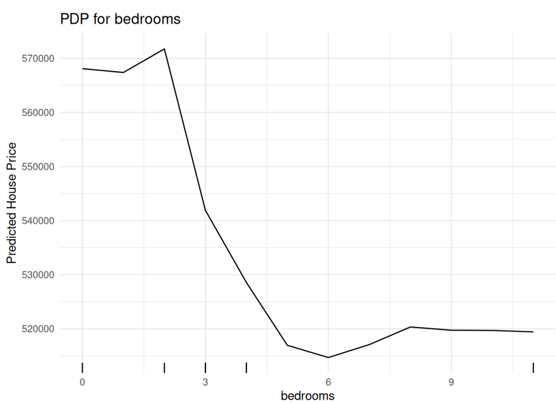
4.2 Model Specification

We used a simple random sample of 5 000 rows is drawn to accelerate PDP evaluation; no rows contain missing values. After that, a Random Forest regressor was implemented with `ranger`, yielding a $R^2 \approx 0.59$, indicating the forest captures a substantial (though not exhaustive) share of price variance.

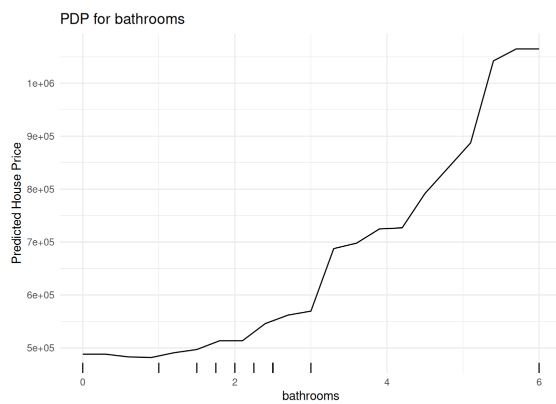
4.3 Partial Dependence Plot

One-dimensional PDPs are computed for the four most salient predictors: `bedrooms`, `bathrooms`, `sqft_living` and `floors`—using the `partial()` function with default grids. Each curve averages over the 5 000-row sample while varying the focal feature.

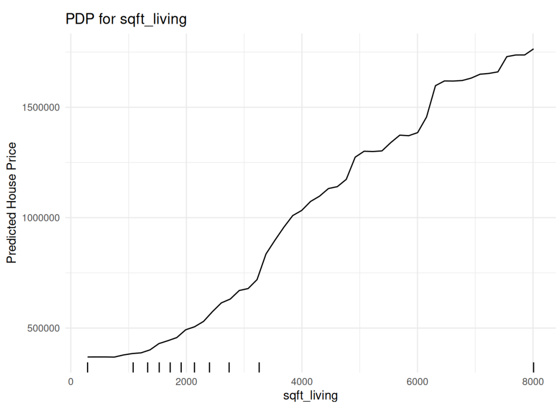
4.4 Results and Interpretation



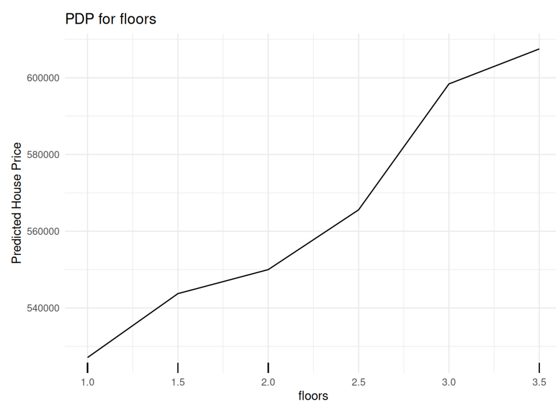
(a) PDP for `bedrooms`



(b) PDP for `bathrooms`



(c) PDP for `sqft_living`



(d) PDP for `floors`

Figure 3: Exercise 3 plots

- **Bedrooms.** Prices rise gently from two to five bedrooms but flatten thereafter, suggesting diminishing returns once a house exceeds the typical family size. Extremely high bedroom counts confer no extra premium.
- **Bathrooms.** A clearer positive slope is visible: each additional bathroom adds appreciable value, even up to 6 bathrooms. Buyers evidently value ample bathroom capacity.
- **Living Area (*sqft_living*).** The relationship is strongly upward. House prices increase exponentially to higher prices. The steepest marginal gains occur beyond 3 000 ft^2, highlighting the premium attached to expansive interiors.
- **Floors.** Moving from single- to multi-storey designs lifts predicted price. Height therefore matters, the higher the more expensive the house is.

5 Conclusion

5.1 Comparative Insights

PDPs expose domain-specific response shapes that would remain opaque in raw permutation-importance tables. For both applications the most influential continuous features show clear non-linearities—linear models would therefore mis-estimate marginal effects. The bike study illustrates how PDPs can separate exogenous influences (weather) from endogenous growth, while the housing study highlights scaling laws (size premiums) and saturation points (bedrooms, floors). In practice, such visual evidence can guide fleet-balancing policies or inform renovation priorities.

5.2 Limitations

- **Independence assumption.** PDPs average over the joint distribution of non-focal features; if these correlate strongly with the focal variable, the curves may include unrealistic combinations and mislead interpretation.
- **Sample representativeness.** The 2-D PDP relied on a 1 000-row slice for tractability, and the house model used a 5 000-row subsample. Although OOB R^2 remained respectable, a full-data refit could alter the surfaces.
- **Global averaging masks heterogeneity.** PDPs show mean effects only; heterogeneous sub-populations or strong interactions beyond two features remain hidden.
- **Moderate explanatory power.** The housing forest captured about 59% of price variance, so its PDPs inherit the model's residual uncertainty.

Overall, Partial Dependence Plots proved to be a fast and intuitive lens for auditing complex ensembles, but richer explanation toolkits and broader data contexts will be essential for production-grade, trustworthy analytics.