Verification Report: PES1UG23AM314

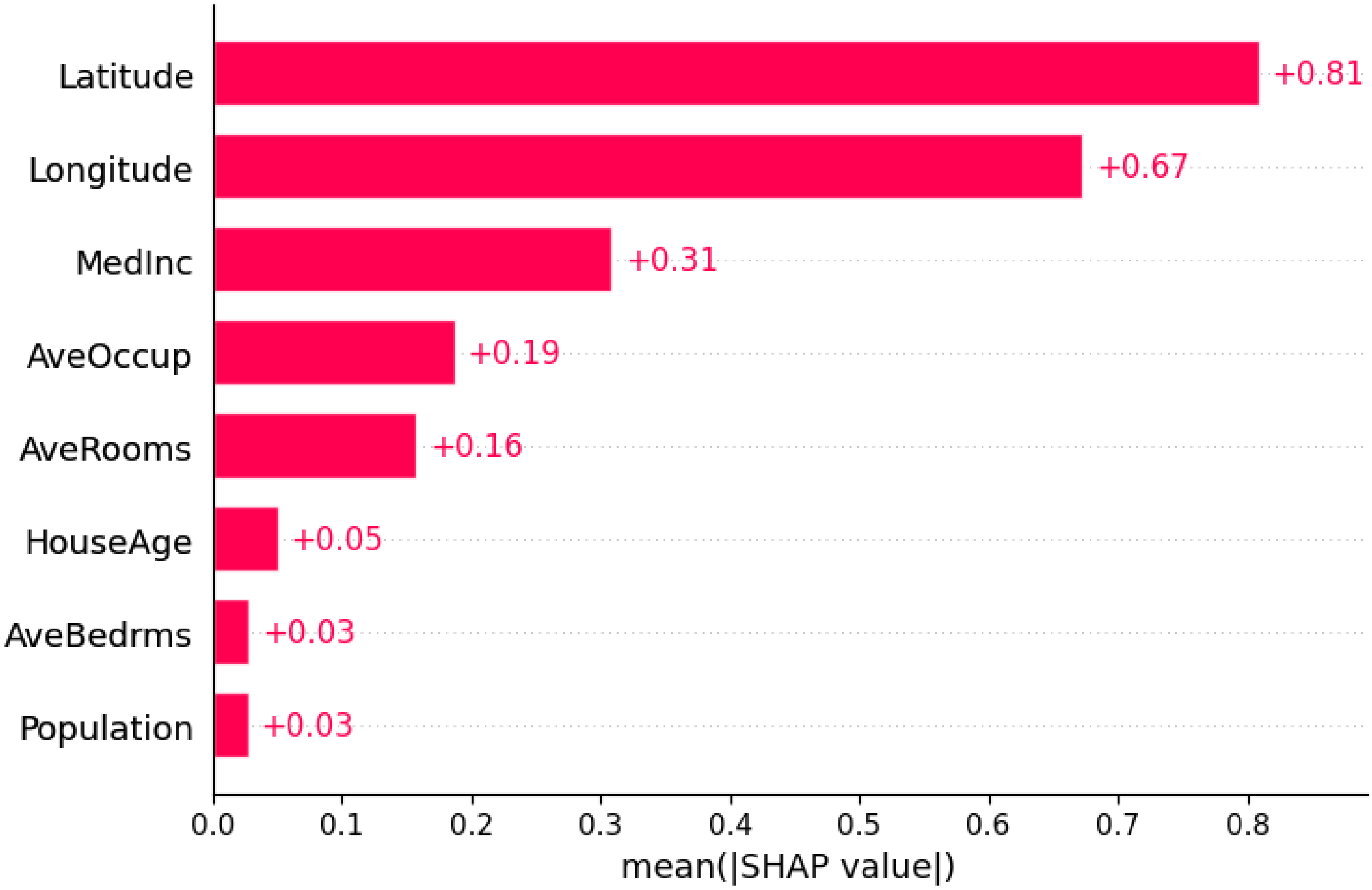
1. Dataset & Model Information

- **Dataset Name**: California Housing Dataset
- **Model Used**: XGBoost Regressor (`XGBRegressor`)
- **Metric Value (Test Set)**:
- **R^2**: 0.8409
- **RMSE**: 0.4566
- **MAPE**: 17.05%

2. SHAP Analysis

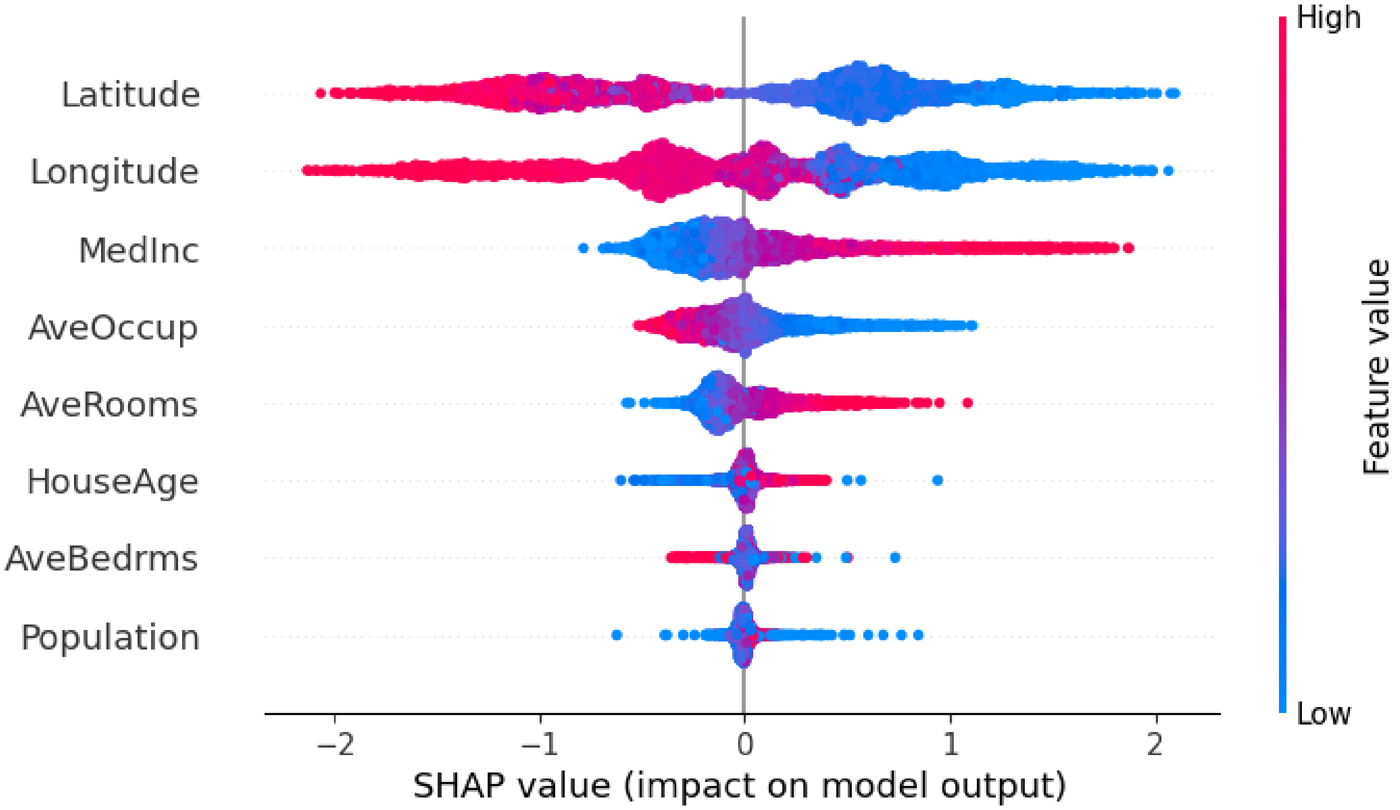
SHAP Bar Plot

Global feature importance based on mean absolute SHAP values.



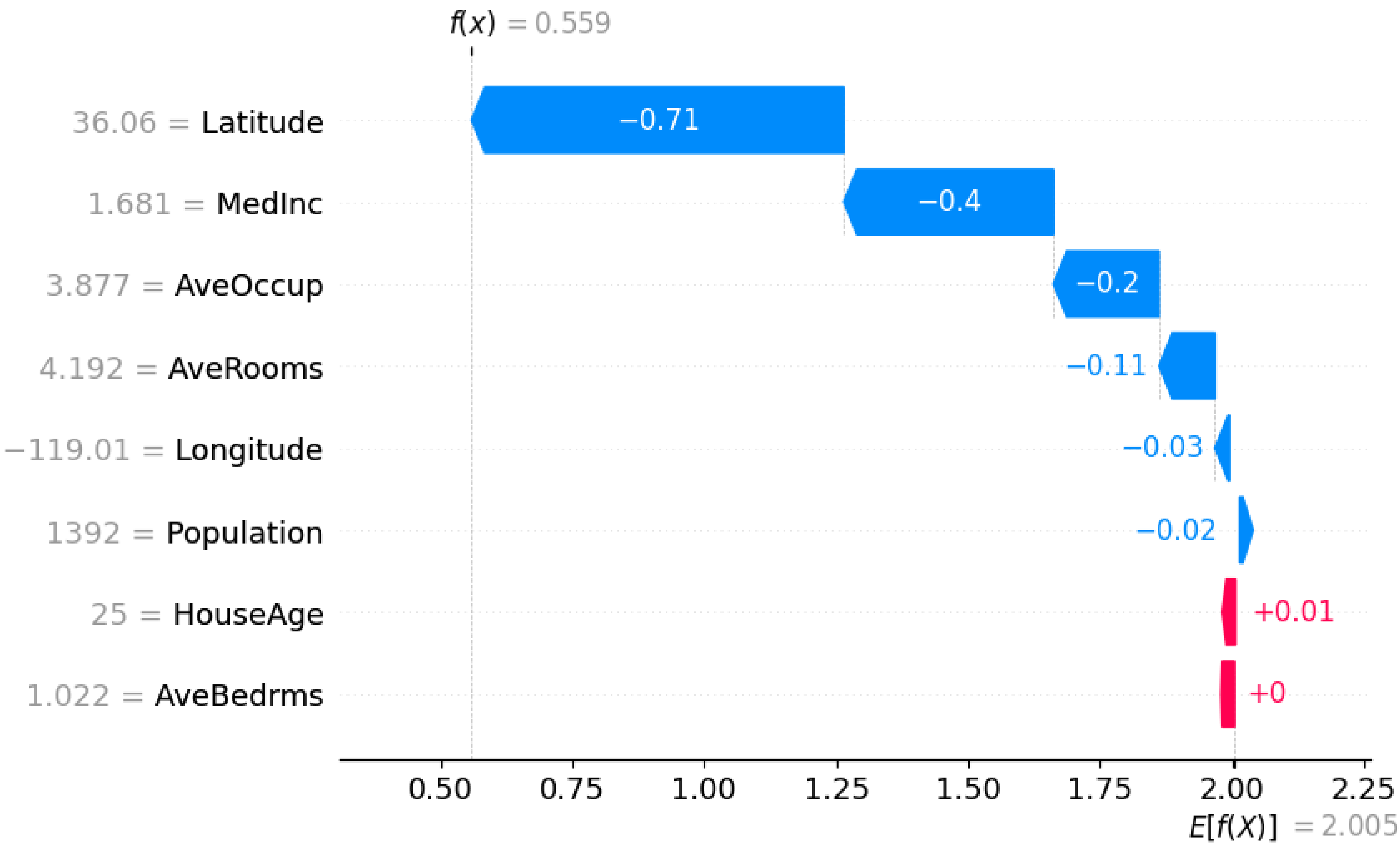
SHAP Summary Plot

Feature impact distribution. Red points = high feature value, Blue points = low feature value.



SHAP Waterfall Plot

Contribution of each feature to the prediction for a single instance.



3. Counterfactual Examples

Counterfactual Example

A generated counterfactual instance that flips the outcome (prediction) to a desired range. *(Based on DiCE generation)*

Feature	Original Value	Counterfactual Value
MedInc	1.68	14.49
HouseAge	25.0	50.3
Prediction	1.0 (approx)	2.04

Actionable Counterfactual

Constraint: Only allow changes in `MedInc` and `HouseAge`.

Feature	Original	Counterfactual	Change
MedInc	1.68	5.69	Increased significantly
HouseAge	25.0	25.0	No change needed
Prediction	1.0	1.22	Increased

Realistic Counterfactual

Constraint: Fix `Latitude` and `Longitude` (location cannot change).

Feature	Original	Counterfactual	Change
MedInc	1.68	6.26	Increased
Population	1392	14174	Increased significantly
Prediction	1.0	1.21	Increased

4. Observations

Observation 1: SHAP Importance ≠ Counterfactual Change

MedInc (Median Income) is the most important feature by SHAP (highest mean |SHAP value|). However, in the actionable counterfactual, changing MedInc alone is sometimes not enough to achieve the desired price increase — HouseAge must also shift. This shows that **SHAP importance** (marginal contribution) **doesn't directly translate to counterfactual actionability**. A feature can be globally important yet require coupled changes with other features to move predictions meaningfully for a specific instance.

Observation 2: Surprising Feature Behaviour

From the SHAP summary plot, **AveOccup** (average occupancy) shows a surprising non-linear behaviour: very low values push predictions **UP** (red dots on positive SHAP side), while moderate-to-high values have near-zero or slightly negative SHAP. Intuitively, fewer occupants per household signals wealthier / less crowded areas, but the effect is disproportionately strong at the low end, suggesting the model captures a **threshold effect** rather than a smooth linear relationship.

Observation 3: Model Bias Observation

The model exhibits **geographic bias** — Latitude and Longitude are among the top SHAP features, meaning the model heavily relies on location to predict house prices. Coastal areas (lower longitude, specific latitude bands near SF/LA) systematically get higher predictions. When we fix Latitude & Longitude in the realistic constraint counterfactual, the model needs **much larger changes** in other features to achieve the same price increase, confirming the model has baked in location-based bias that may disadvantage inland communities regardless of other housing quality factors.