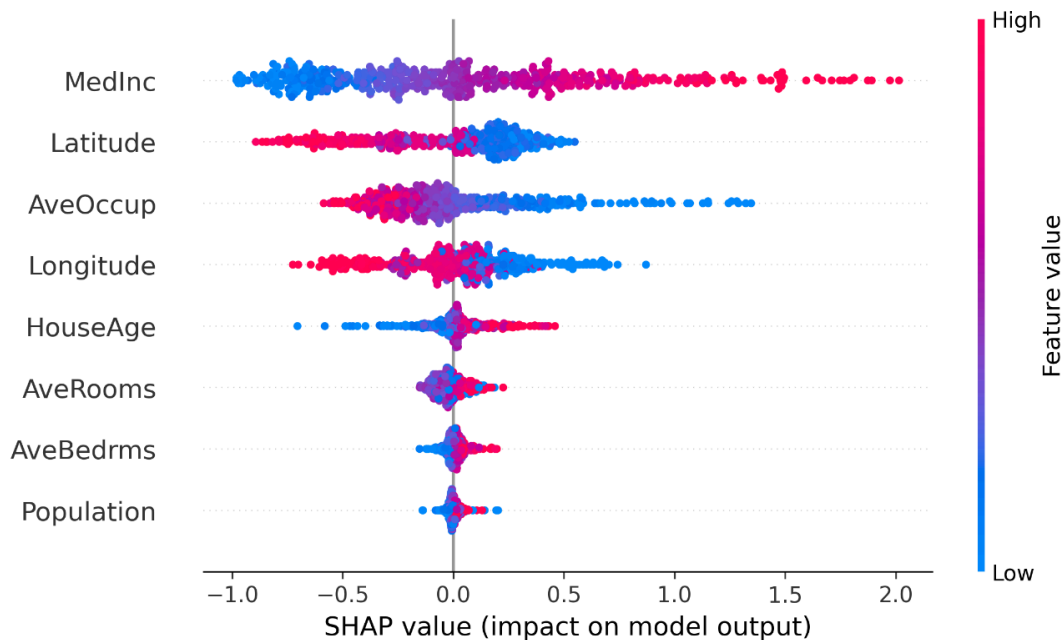


Interprétation des importances globales des features à l'aide de SHAP



Le graphique dans **shap_summary_plot.png** présente la distribution des valeurs SHAP pour chaque variable, où l'ordre vertical des variables est déterminé par leur importance globale. Les variables situées en haut du graphique ont un impact plus significatif sur les prédictions du modèle. L'étalement horizontal des points pour chaque variable montre son niveau d'influence : plus les points sont dispersés, plus la variable a un impact notable sur les prédictions.

MedInc (Revenu médian) : La variable la plus influente, avec des revenus élevés (points rouges) augmentant le prix des logements et des revenus faibles (points bleus) ayant l'effet inverse.

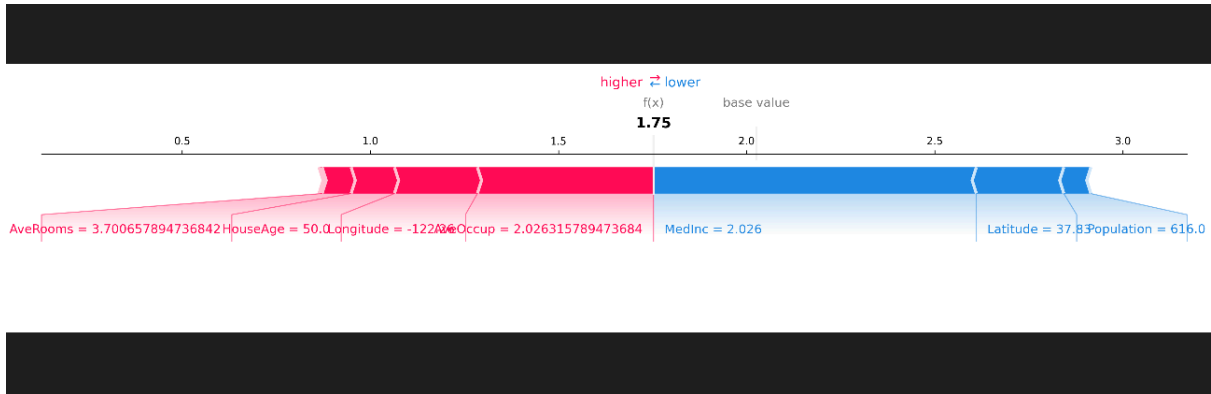
Latitude : La latitude présente également un impact notable, bien que moins marqué que celui de MedInc.

Longitude : L'impact de la longitude est similaire à celui de la latitude.

AveOccup (Occupation moyenne des logements) : Cette variable montre une influence principalement négative. En effet, les points rouges (indiquant une occupation élevée) sont majoritairement situés à gauche de zéro, suggérant que des niveaux élevés d'occupation des logements sont associés à des prix plus bas.

Autres variables : Les variables **HouseAge**, **AveRooms**, **AveBedrms**, et **Population** semblent avoir un impact relativement faible sur les prédictions du modèle par rapport aux variables précédemment mentionnées.

Analyse de l'impact local pour des exemples individuels



Le graphique ci-dessous, **un force plot SHAP** (présenté dans shap_plot_sample_42.png) , montre la décomposition de la prédiction pour un exemple individuel. La **base value**, représentant la prédiction moyenne du modèle, est de 1.75.

Les contributions des différentes variables sont visualisées par des flèches : les flèches rouges indiquent une contribution positive (augmentation de la prédiction) et les flèches bleues une contribution négative (diminution de la prédiction). La longueur des flèches reflète l'amplitude de chaque contribution.

Les variables ayant le plus contribué positivement à la prédiction sont AveRooms, Longitude, et AveOccup, tandis que celles ayant une contribution négative sont MedInc, HouseAge, et Population.