Model Interpretability Report

Model Information:

Model Type: LinearRegression

Model Parameters: {'copy_X': True, 'fit_intercept': True, 'n_jobs': None, 'positive': False}

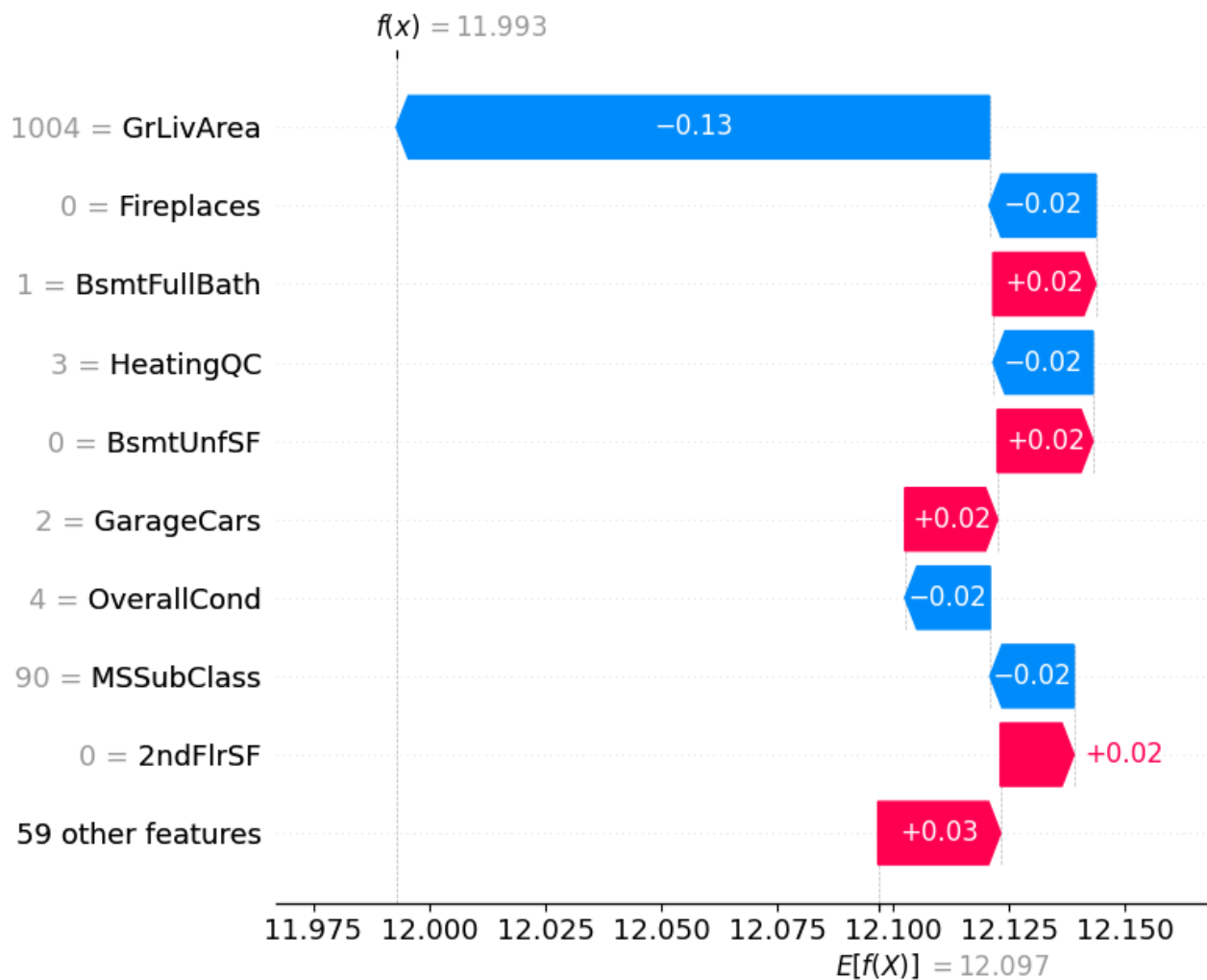**Data Description:**

Number of Samples: 292

Number of Features: 68

Feature Names: MSSubClass, MSZoning, LotFrontage, LotArea, Street, LotShape, LandContour, Utilities, LotConfig, LandSlope, Neighborhood, Condition1, Condition2, HouseStyle, OverallQual, OverallCond, YearBuilt, YearRemodAdd, RoofStyle, RoofMatl, Exterior1st, MasVnrType, MasVnrArea, ExterQual, ExterCond, Foundation, BsmtQual, BsmtCond, BsmtExposure, BsmtFinType1, BsmtFinSF1, BsmtFinType2, BsmtUnfSF, TotalBsmtSF, Heating, HeatingQC, CentralAir, Electrical, 2ndFlrSF, LowQualFinSF, GrLivArea, BsmtFullBath, BsmtHalfBath, FullBath, HalfBath, BedroomAbvGr, KitchenAbvGr, KitchenQual, Functional, Fireplaces, FireplaceQu, GarageType, GarageFinish, GarageCars, GarageQual, GarageCond, PavedDrive, WoodDeckSF, OpenPorchSF, EnclosedPorch, 3SsnPorch, ScreenPorch, PoolArea, MiscVal, MoSold, YrSold, SaleType, SaleCondition

This dataset was used to generate SHAP values, which help interpret the model predictions.

Waterfall Plot for Single Prediction

This plot illustrates how each feature contributes to a single prediction. Positive SHAP values push the prediction higher, while negative values decrease it.

$f(x) = 11.993$

1004 = GrLivArea	−0.13
0 = Fireplaces	−0.02
1 = BsmtFullBath	+0.02
3 = HeatingQC	−0.02
0 = BsmtUnfSF	+0.02
2 = GarageCars	+0.02
4 = OverallCond	−0.02
90 = MSSubClass	−0.02
0 = 2ndFlrSF	+0.02
59 other features	+0.03

11.975 12.000 12.025 12.050 12.075 12.100 12.125 12.150

$E[f(X)] = 12.097$

Feature Importance Bar Plot

This bar plot ranks the features by their average SHAP values, indicating the overall importance of each feature in making predictions.
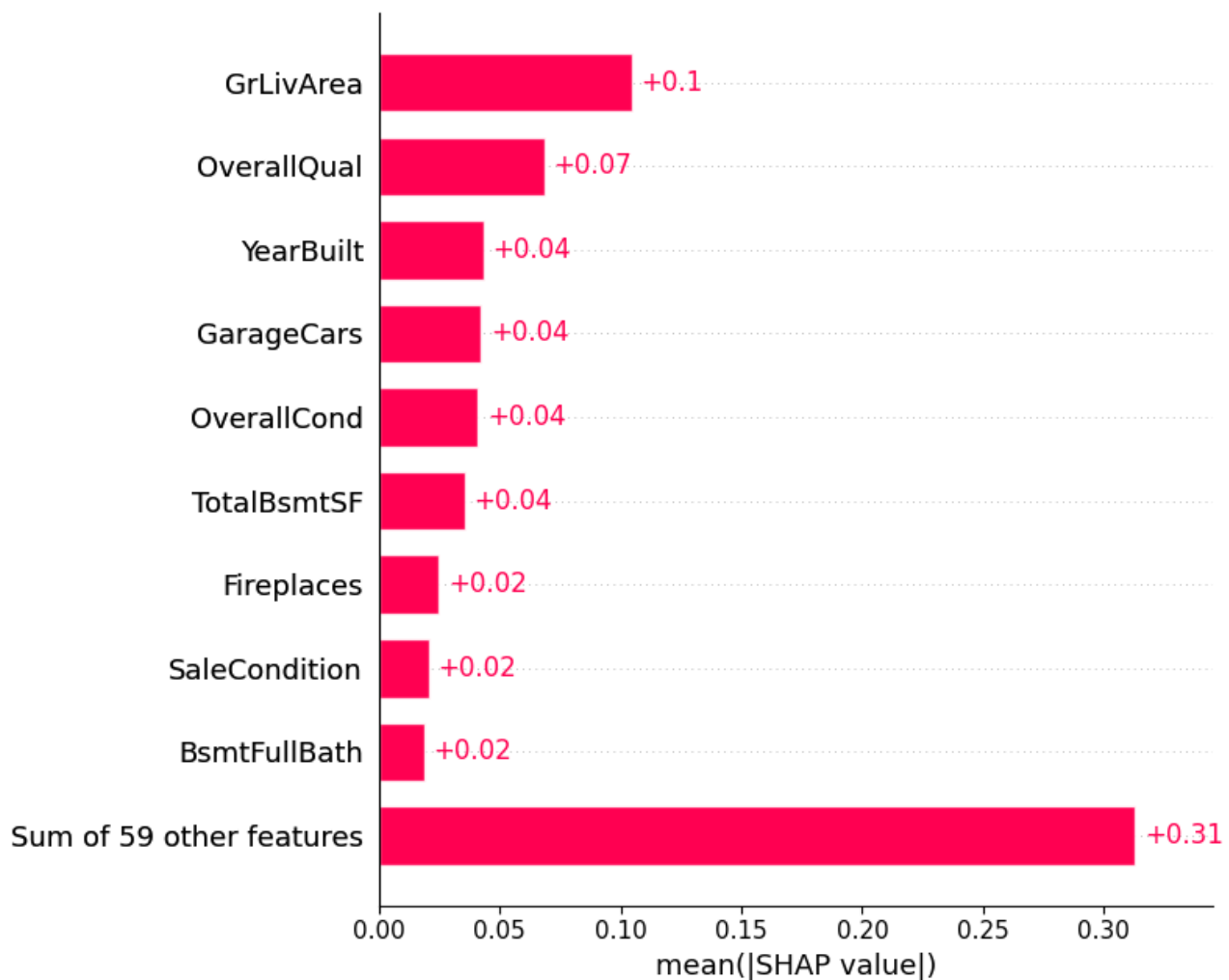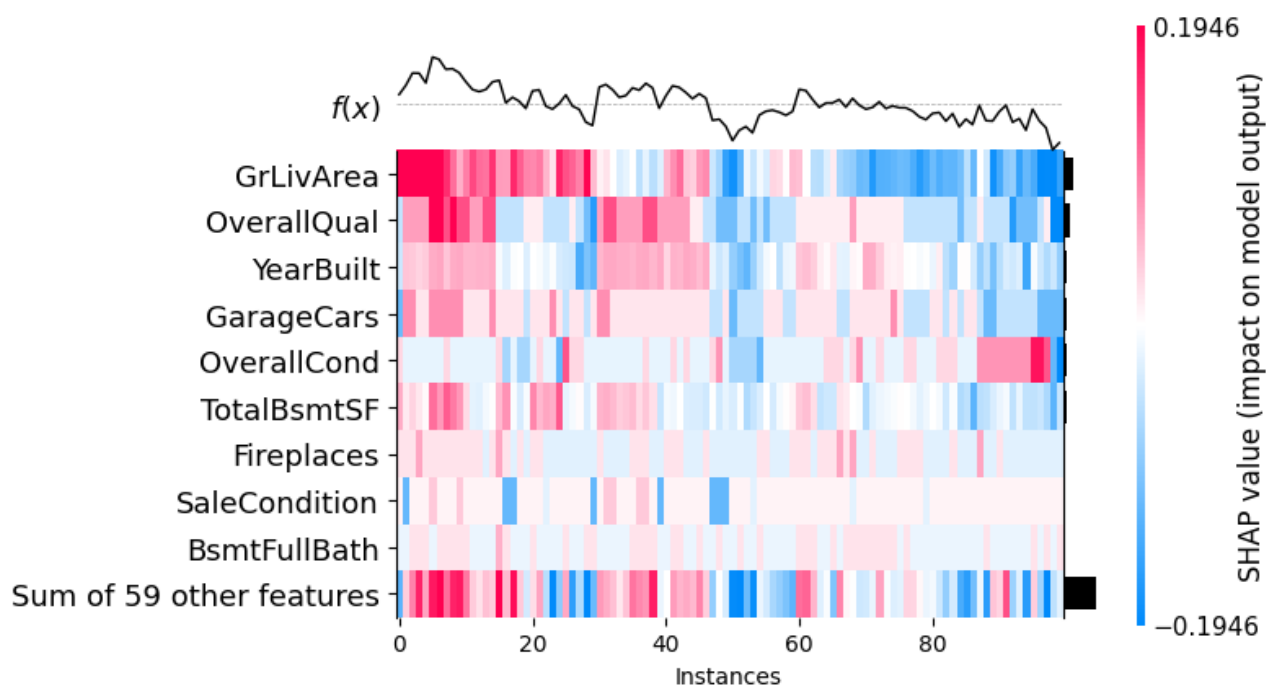
Beeswarm Plot

The beeswarm plot is designed to display an information-dense summary of how the topfeatures in a dataset impact the model output.
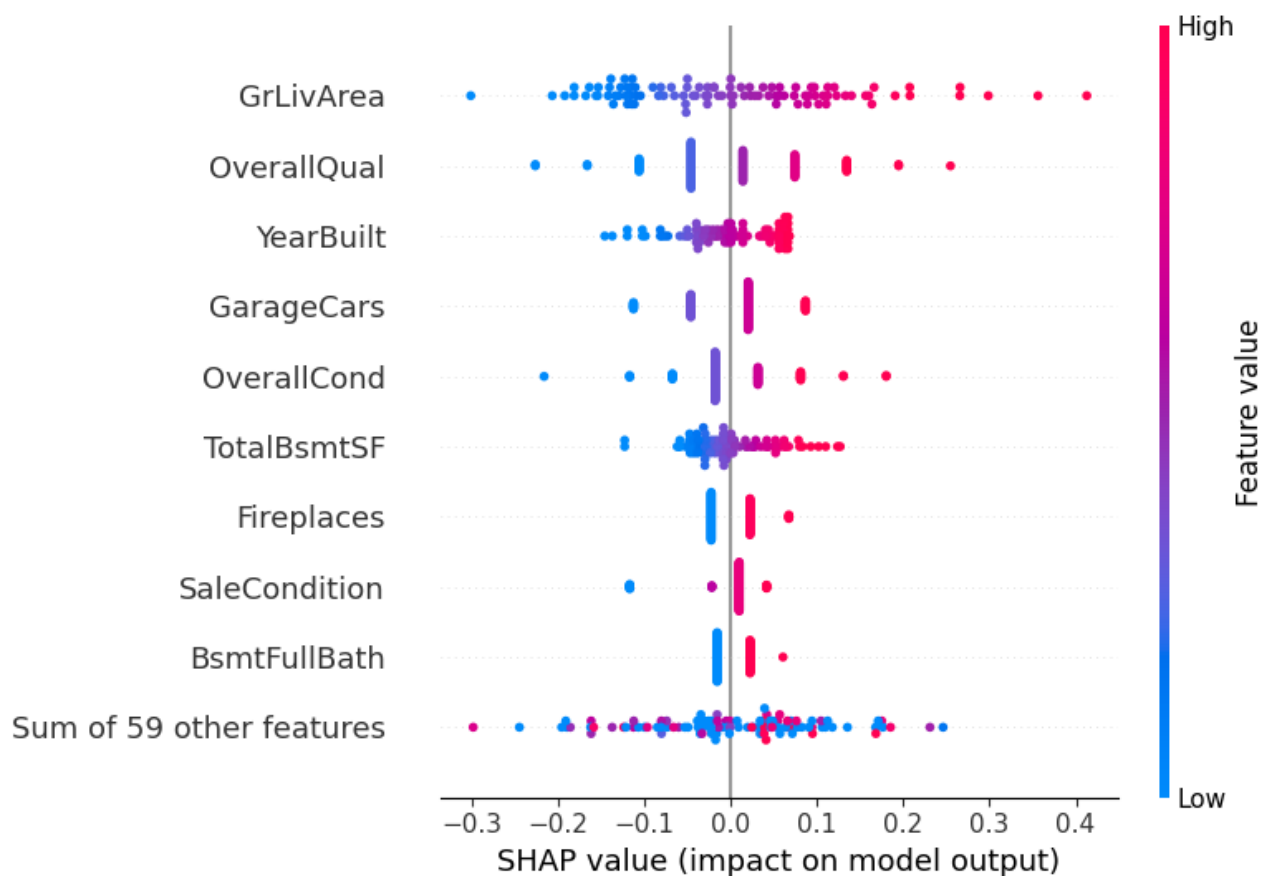
Heatmap Plot

The heatmap plot function creates a plot with the instances on the x-axis, the modelinputs on the y-axis, and the SHAP values encoded on a color scale.

Interpretation Summary:

The SHAP analysis provides both global and local interpretability for the model. Globally, we understand which features are most influential across the entire dataset, while locally, the waterfall plot allows us to dissect individual predictions.