Real Estate Price Prediction with Elastic-net Regression

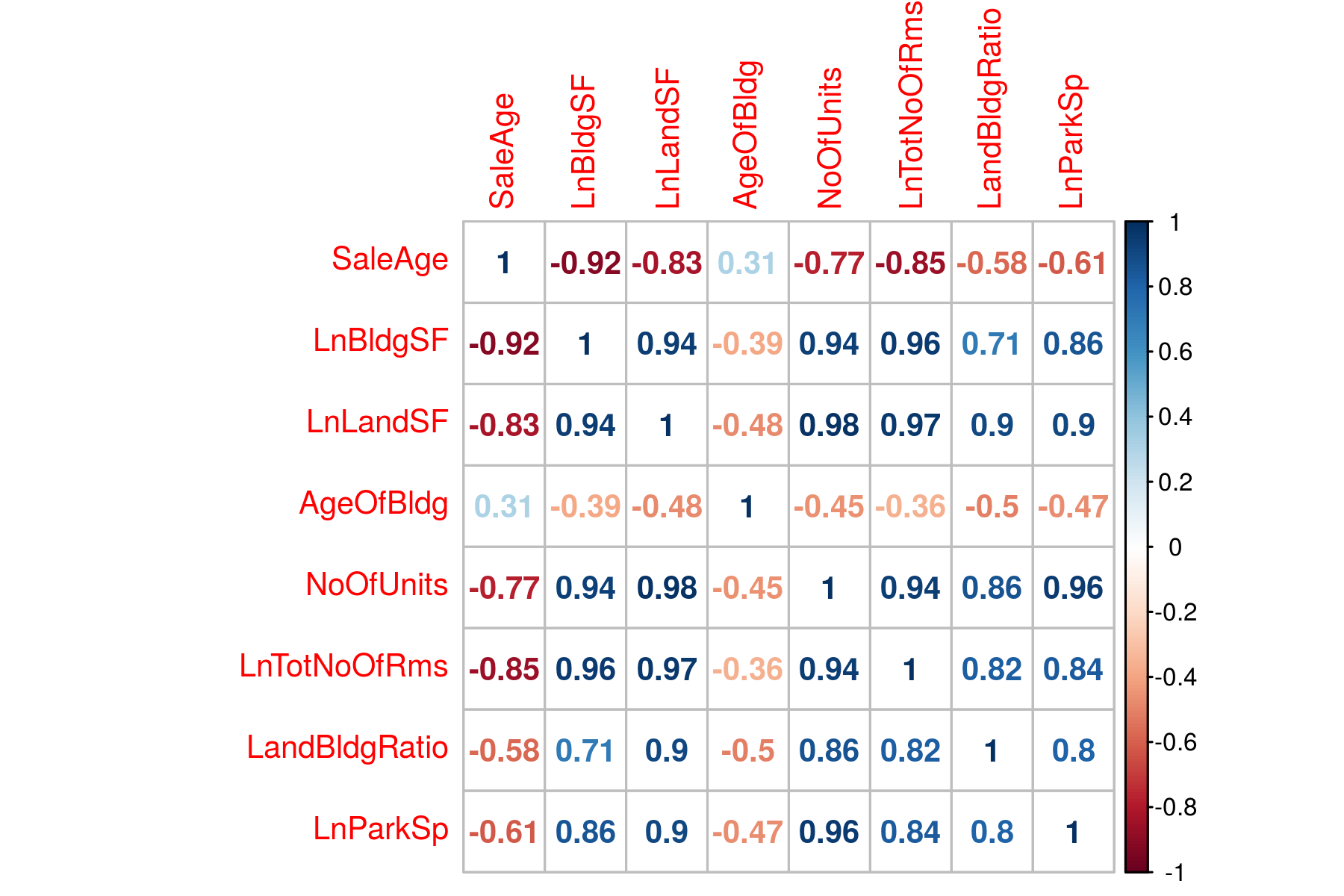
Input Data file: data/AptDataSet.csv

## Basic summary statistics

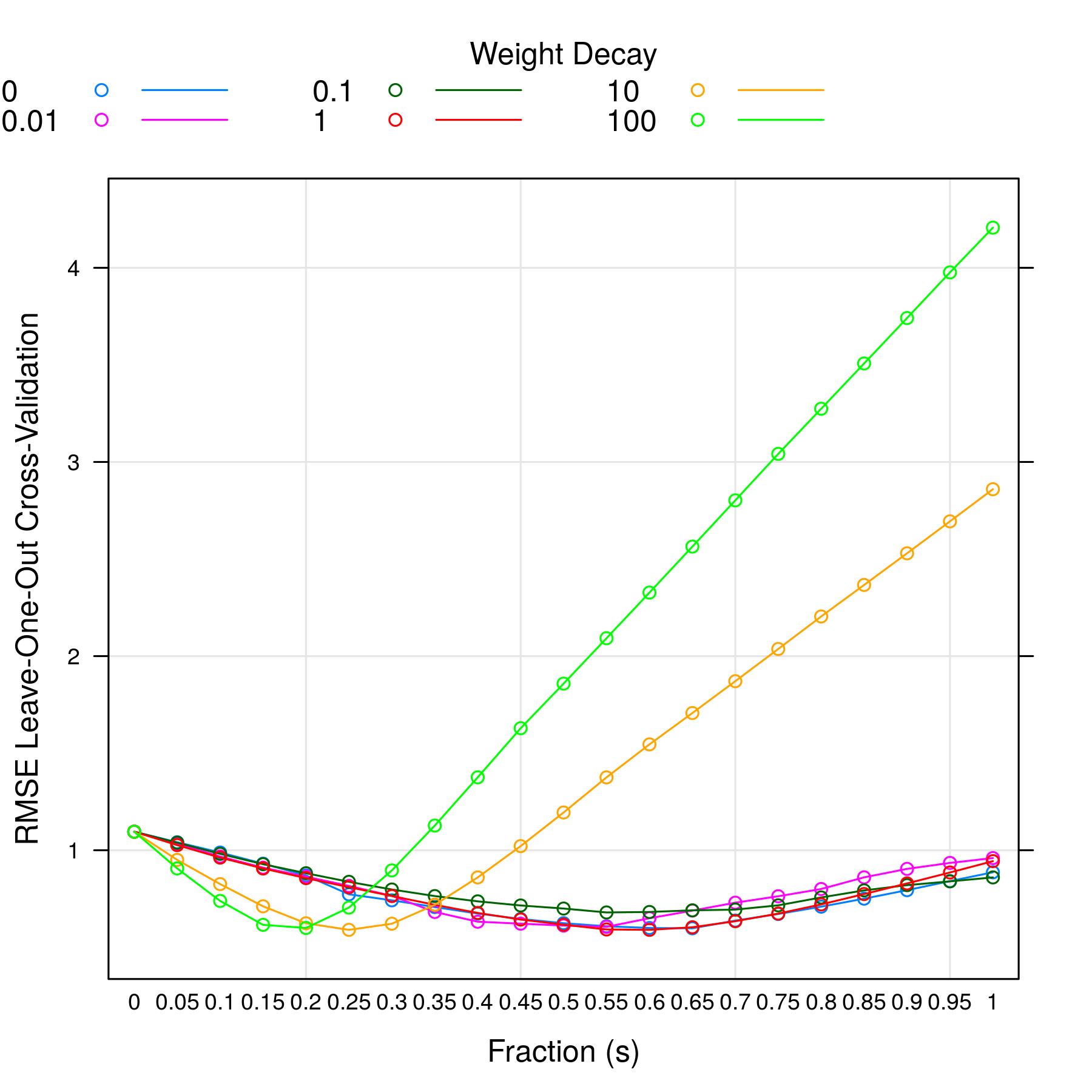
|  | **N** | **Mean** | **Std Dev** | **Min** | **Max** |
| --- | --- | --- | --- | --- | --- |
| LnSalePrice | 006 | 014 | 000 | 014 | 015 |
| SaleAge | 006 | 645 | 387 | 009 | 1100 |
| LnBldgSF | 006 | 010 | 000 | 010 | 011 |
| LnLandSF | 006 | 011 | 000 | 011 | 012 |
| AgeOfBldg | 006 | 041 | 007 | 034 | 049 |
| NoOfUnits | 006 | 077 | 026 | 050 | 114 |
| LnTotNoOfRms | 006 | 005 | 000 | 004 | 005 |
| LandBldgRatio | 006 | 002 | 000 | 001 | 002 |
| LnParkSp | 006 | 004 | 000 | 004 | 005 |

NOTE - No summary statistics are provided for categorical variables.

## Correlations Between Predictors

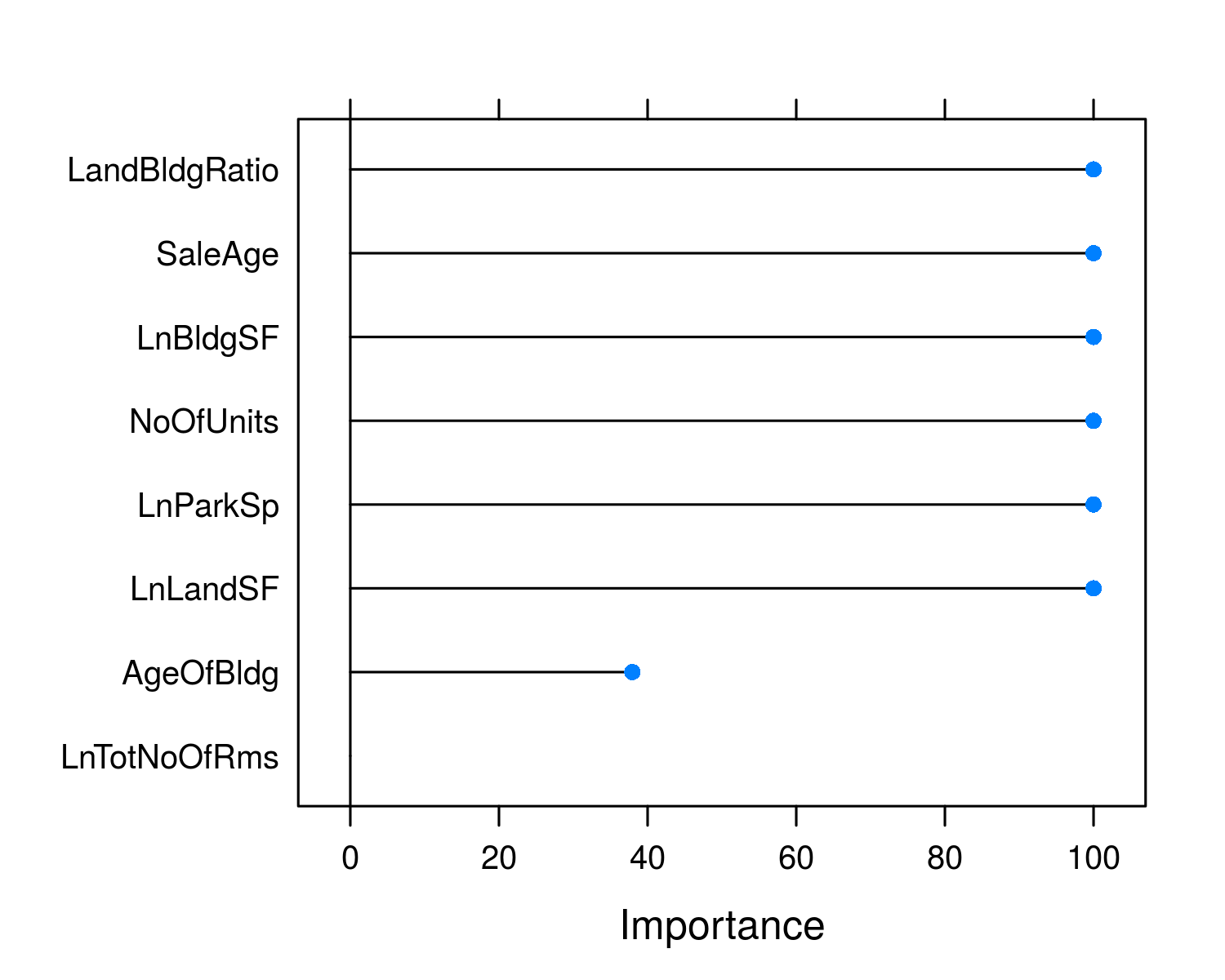


## Tuning Parameter Selection Using LOOCV



From above plot, lambda =1.00 and s =0.60 gives the minimum RMSE model.

## Variable Importance



## Standardized Model Coefficients

| **Variable** | **Estimate** |
| --- | --- |
| SaleAge | -0.248 |
| LnBldgSF | 0.437 |
| LnLandSF | 0.238 |
| AgeOfBldg | 0.000 |
| NoOfUnits | 0.207 |
| LnTotNoOfRms | 0.154 |
| LandBldgRatio | 0.000 |
| LnParkSp | 0.157 |

NOTE std. errors are calculated using bootstrapping which is the only way to determine coef. errors for a penalized regression. But the errors should be only used for reference. It is yet unclear how meaningful the std. errors are in penalized regression.

## Model Prediction

| **Predicted Value** | **Prediction Error** | **R2** |
| --- | --- | --- |
| 16.201 | 0.590 | 0.608 |