Real Estate Price Prediction with Elastic-net Regression

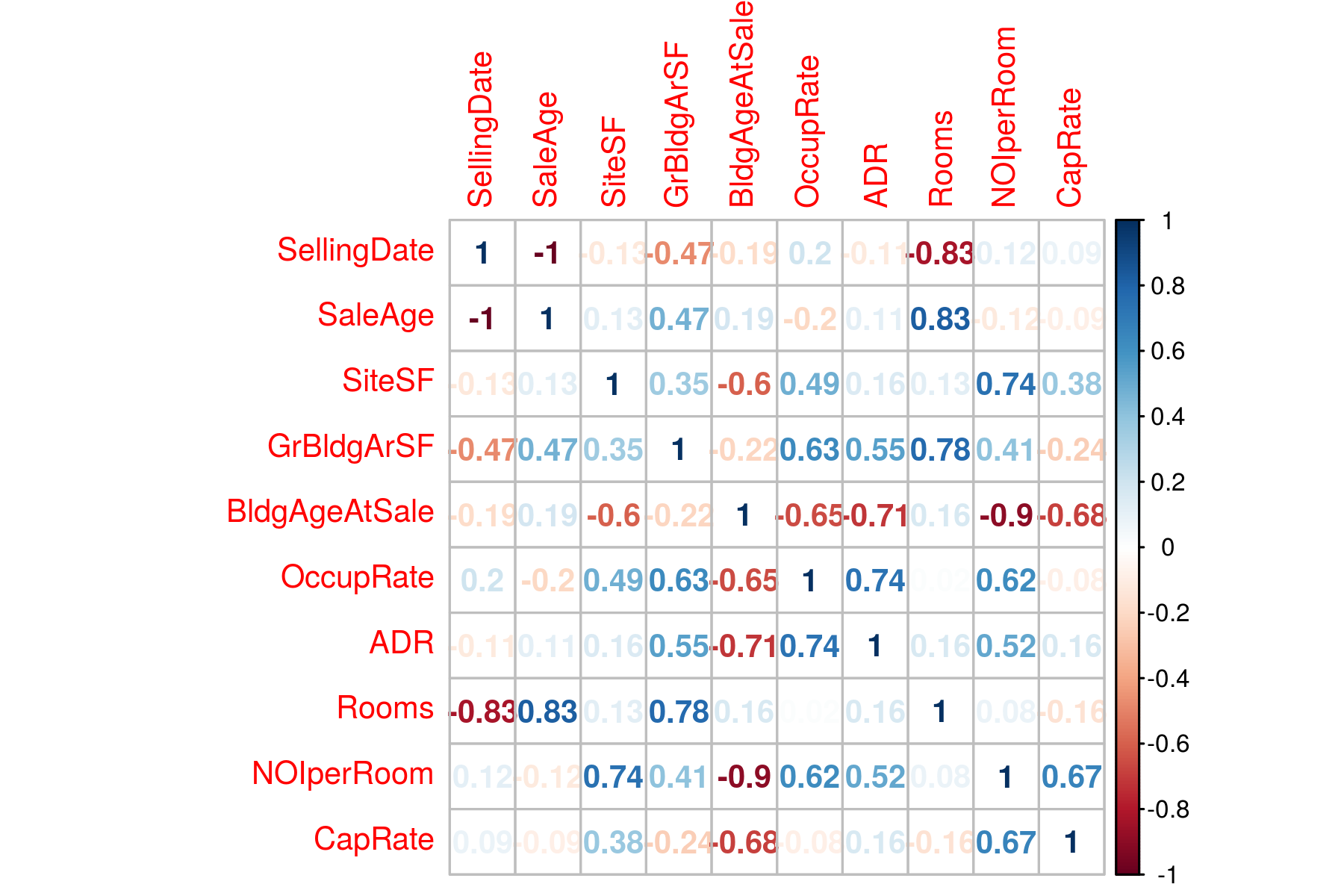
Input Data file: data/Motel\_mod.csv

## Basic summary statistics

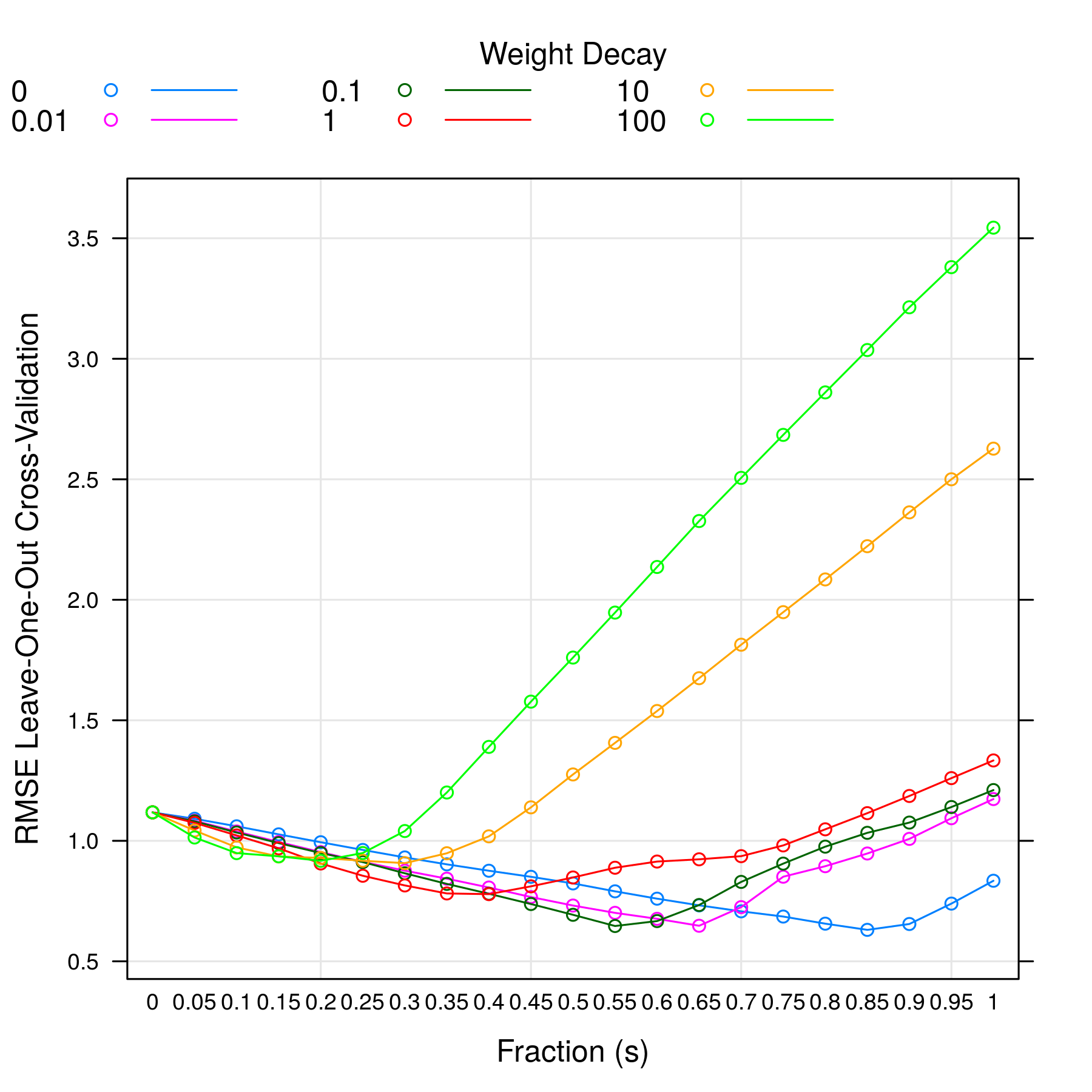
|  | **N** | **Mean** | **Std Dev** | **Min** | **Max** |
| --- | --- | --- | --- | --- | --- |
| SellingPrice | 005 | 3100300 | 2164038 | 1010000 | 6427500 |
| SellingDate | 005 | 42405 | 627 | 41439 | 42866 |
| SaleAge | 005 | 644 | 627 | 183 | 1610 |
| SiteSF | 005 | 77439 | 23353 | 47306 | 111949 |
| GrBldgArSF | 005 | 28370 | 10594 | 14400 | 43041 |
| BldgAgeAtSale | 005 | 023 | 007 | 016 | 034 |
| OccupRate | 005 | 058 | 010 | 048 | 073 |
| ADR | 005 | 068 | 020 | 052 | 103 |
| Rooms | 005 | 064 | 020 | 035 | 087 |
| NOIperRoom | 005 | 4727 | 2987 | 1195 | 8224 |
| CapRate | 005 | 000 | 000 | 000 | 000 |

NOTE - No summary statistics are provided for categorical variables.

## Correlations Between Predictors

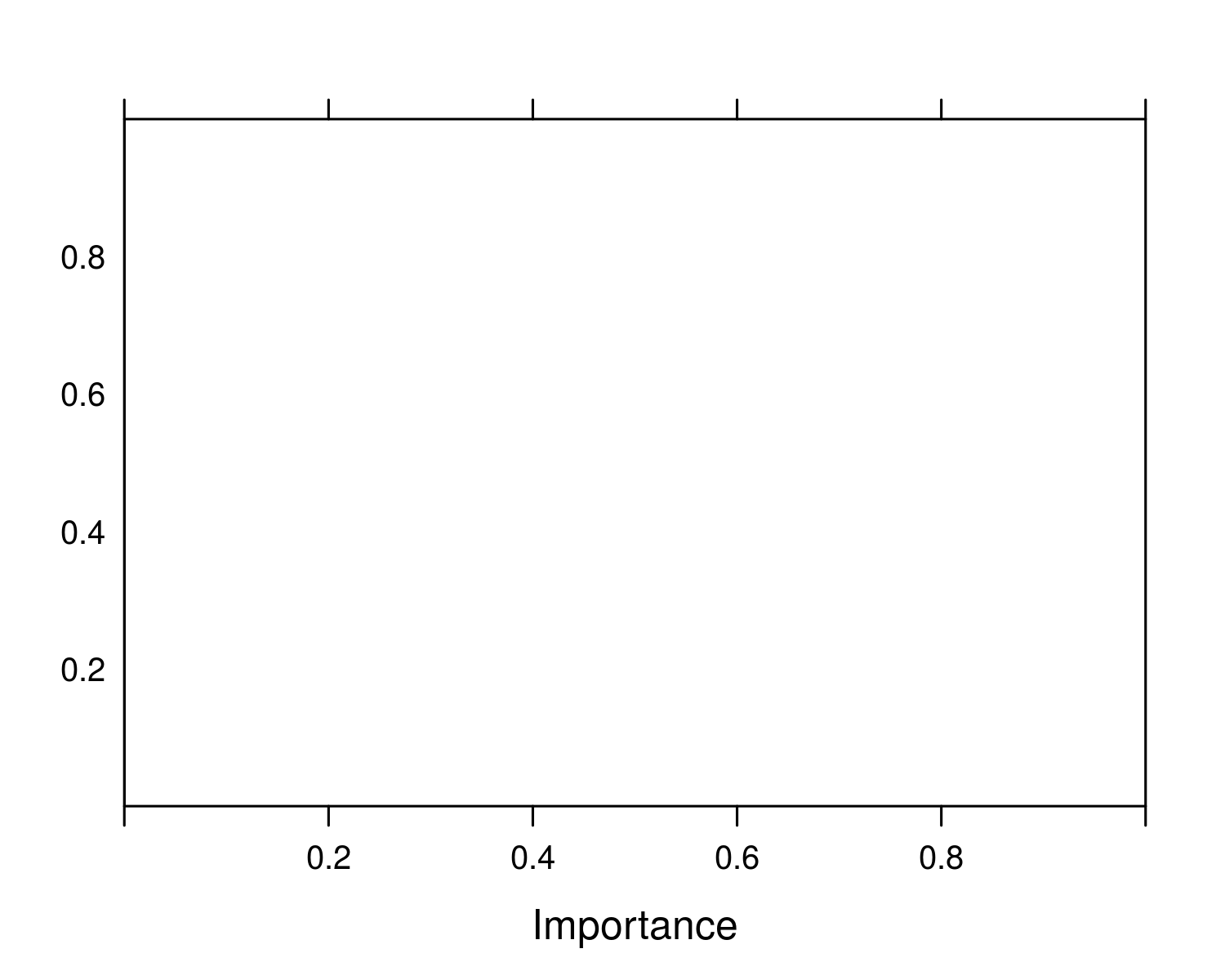


## Tuning Parameter Selection Using LOOCV



From above plot, lambda =0.00 and s =0.85 gives the minimum RMSE model.

## Variable Importance



## Standardized Model Coefficients

| **Variable** | **Estimate** |
| --- | --- |
| SiteSF | 0.000 |
| GrBldgArSF | 0.519 |
| OccupRate | 0.000 |
| ADR | 0.170 |
| NOIperRoom | 0.375 |

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** |
| --- | --- | --- |
| 2241181.386 | 0.631 | 0.562 |