Real Estate Price Prediction with Elastic-net Regression

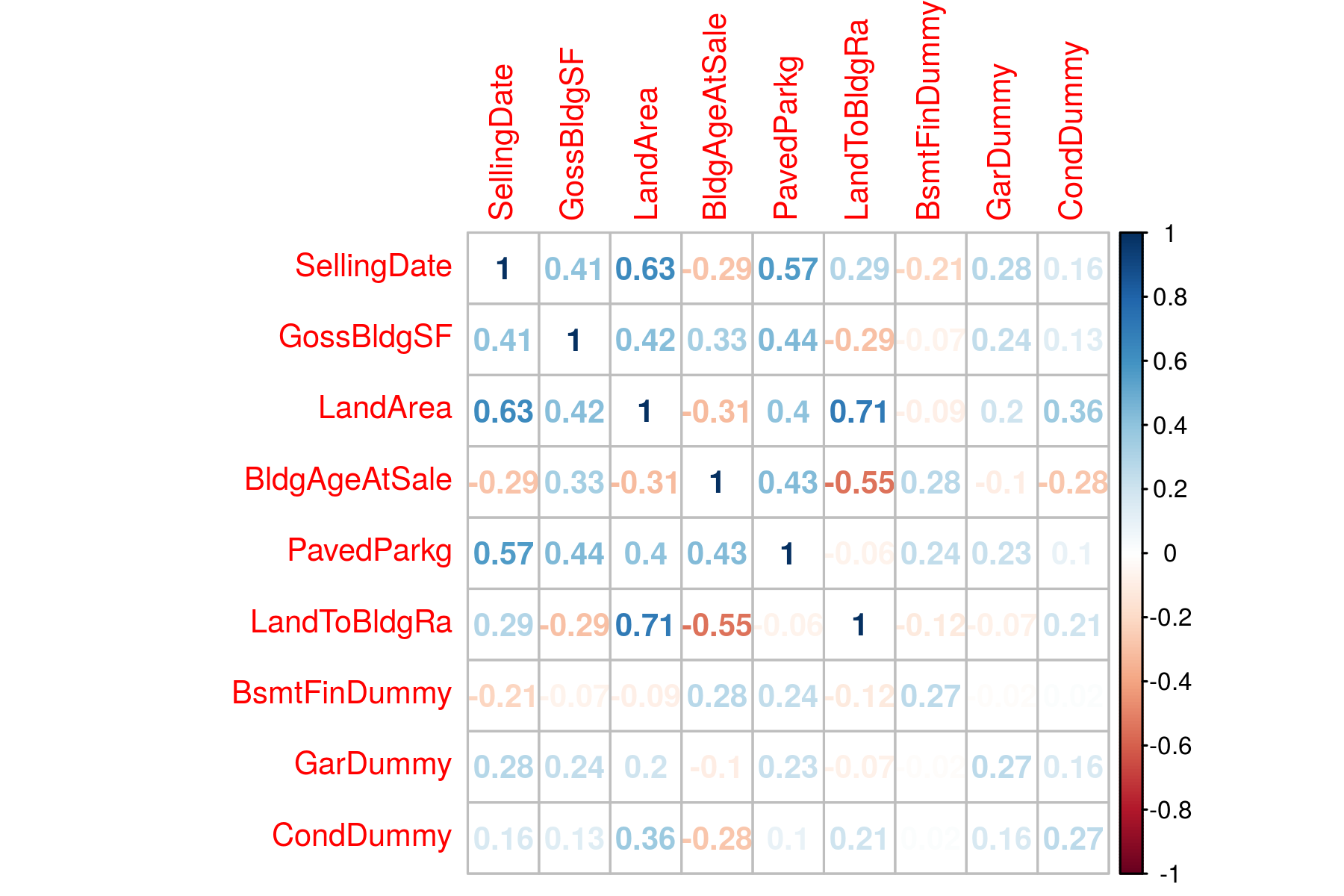
Input Data file: data/Church\_mod.csv

## Basic summary statistics

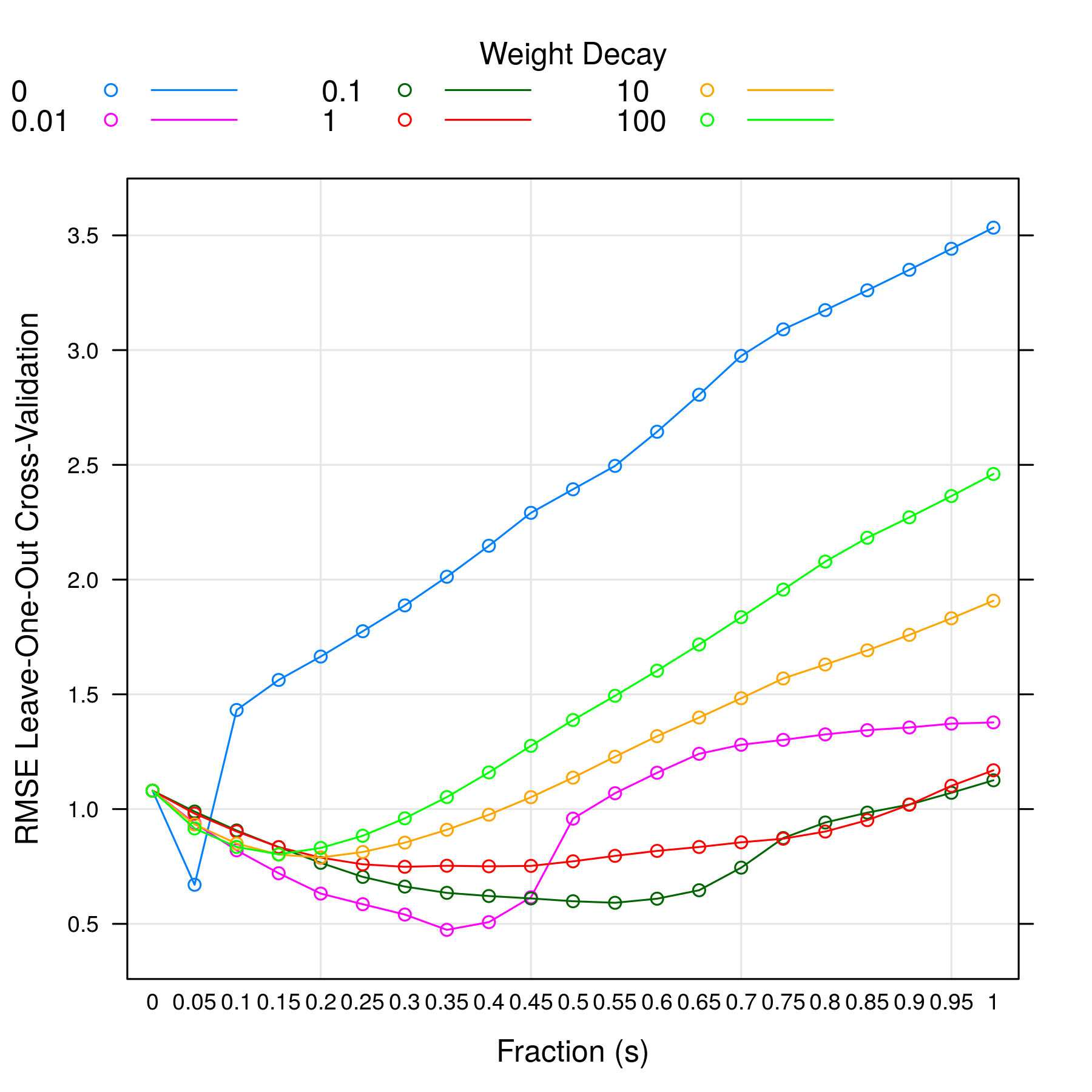
|  | **N** | **Mean** | **Std Dev** | **Min** | **Max** |
| --- | --- | --- | --- | --- | --- |
| SellingPrice | 007 | 211421 | 111145 | 65000 | 344950 |
| SellingDate | 007 | 42296 | 474 | 41547 | 42724 |
| GossBldgSF | 007 | 6701 | 3910 | 2206 | 13363 |
| BsmtFinDummy | 007 | 000 | 000 | 000 | 001 |
| LandArea | 007 | 23102 | 15098 | 6500 | 51357 |
| BldgAgeAtSale | 007 | 061 | 031 | 030 | 109 |
| PavedParkg | 007 | 369 | 528 | 000 | 1363 |
| GarDummy | 007 | 000 | 000 | 000 | 001 |
| CondDummy | 007 | 000 | 000 | 000 | 001 |
| LandToBldgRa | 007 | 003 | 002 | 001 | 010 |

NOTE - No summary statistics are provided for categorical variables.

## Correlations Between Predictors

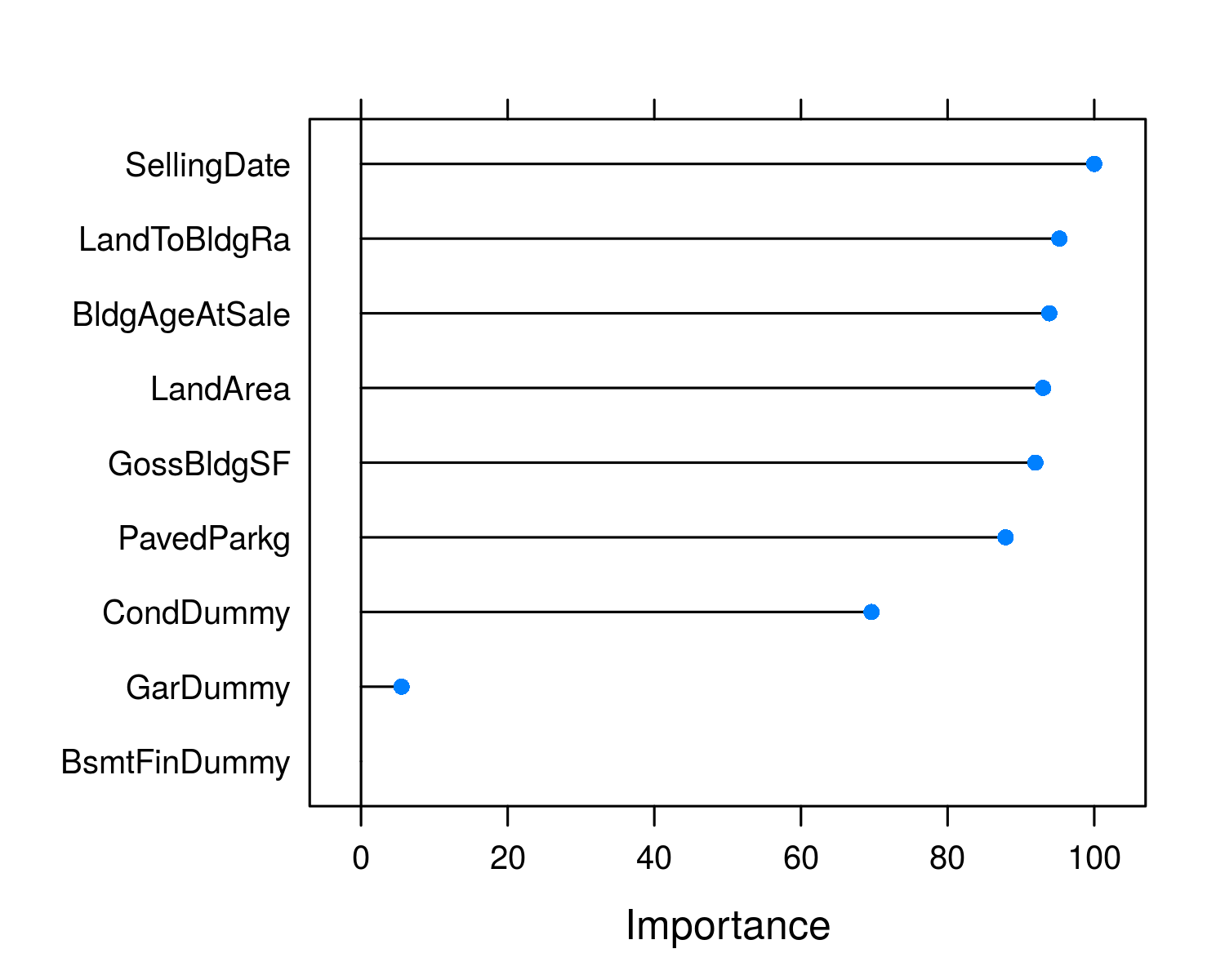


## Tuning Parameter Selection Using LOOCV



From above plot, lambda =0.01 and s =0.35 gives the minimum RMSE model.

## Variable Importance



## Standardized Model Coefficients

| **Variable** | **Estimate** |
| --- | --- |
| SellingDate | 0.000 |
| GossBldgSF | 0.104 |
| LandArea | 0.410 |
| BldgAgeAtSale | 0.000 |
| PavedParkg | 0.000 |
| LandToBldgRa | 0.000 |
| BsmtFinDummy | 0.000 |
| GarDummy | 0.000 |
| CondDummy | 0.719 |

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** |
| --- | --- | --- |
| 317625.424 | 0.474 | 0.779 |