Real Estate Price Prediction with Elastic-net Regression

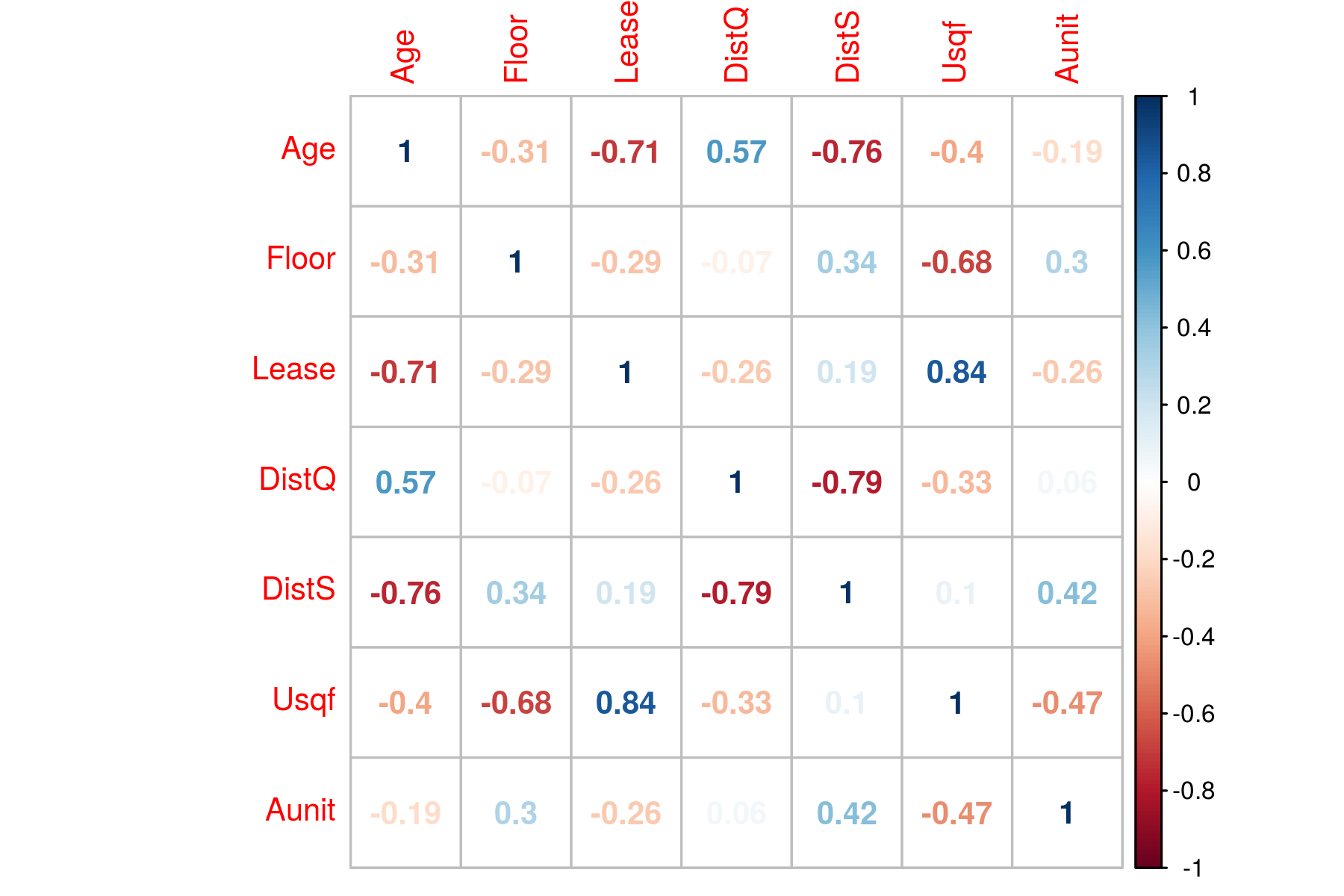
Input Data file: data/Office\_revised.csv

## Basic summary statistics

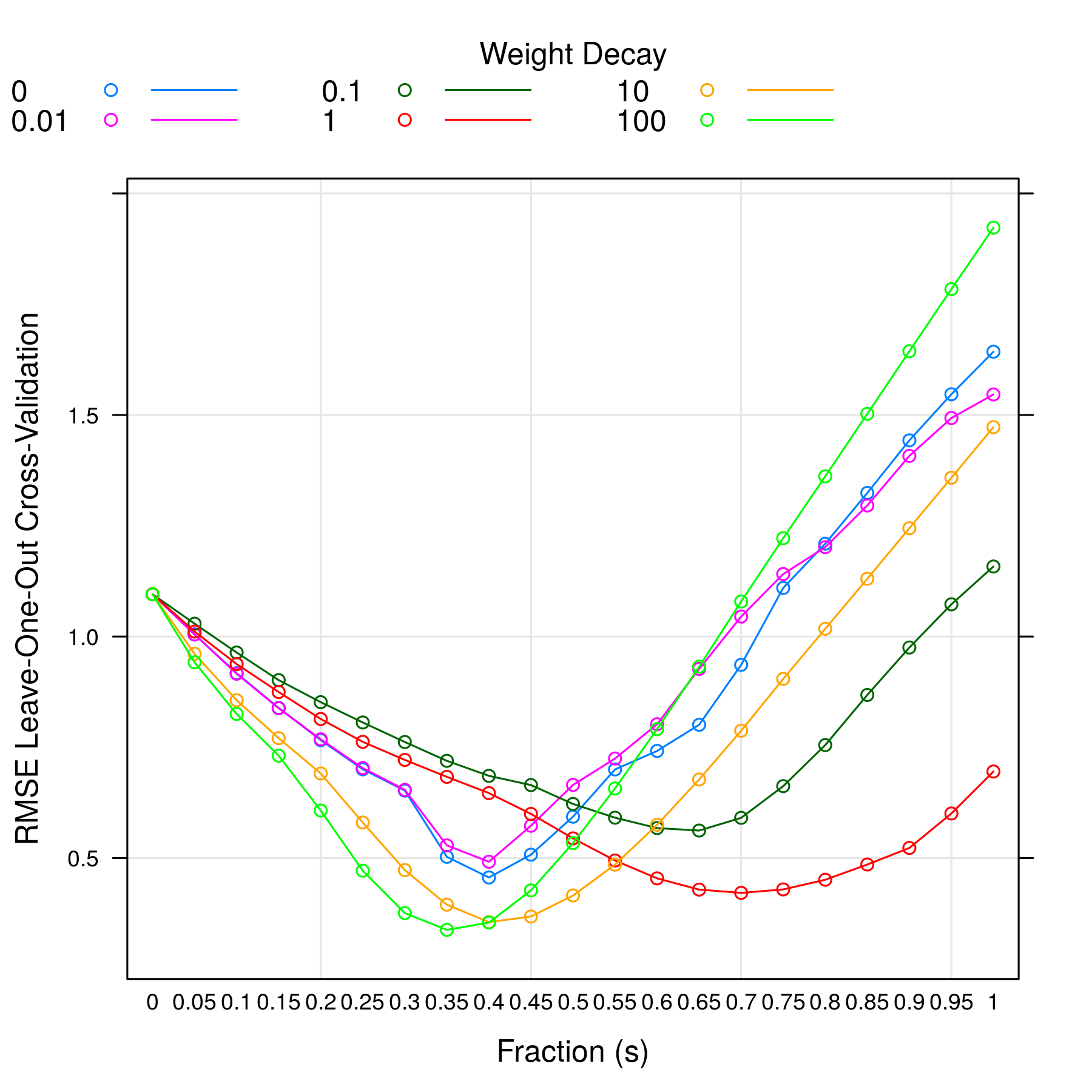
|  | **N** | **Mean** | **Std Dev** | **Min** | **Max** |
| --- | --- | --- | --- | --- | --- |
| Rent | 006 | 004 | 001 | 002 | 007 |
| Age | 006 | 040 | 037 | 006 | 080 |
| Floor | 006 | 004 | 003 | 001 | 012 |
| Lease | 006 | 001 | 001 | 000 | 005 |
| DistQ | 006 | 000 | 000 | 000 | 000 |
| DistS | 006 | 000 | 000 | 000 | 000 |
| Usqf | 006 | 1340 | 637 | 530 | 2101 |
| Aunit | 006 | 003 | 002 | 001 | 007 |

NOTE - No summary statistics are provided for categorical variables.

## Correlations Between Predictors

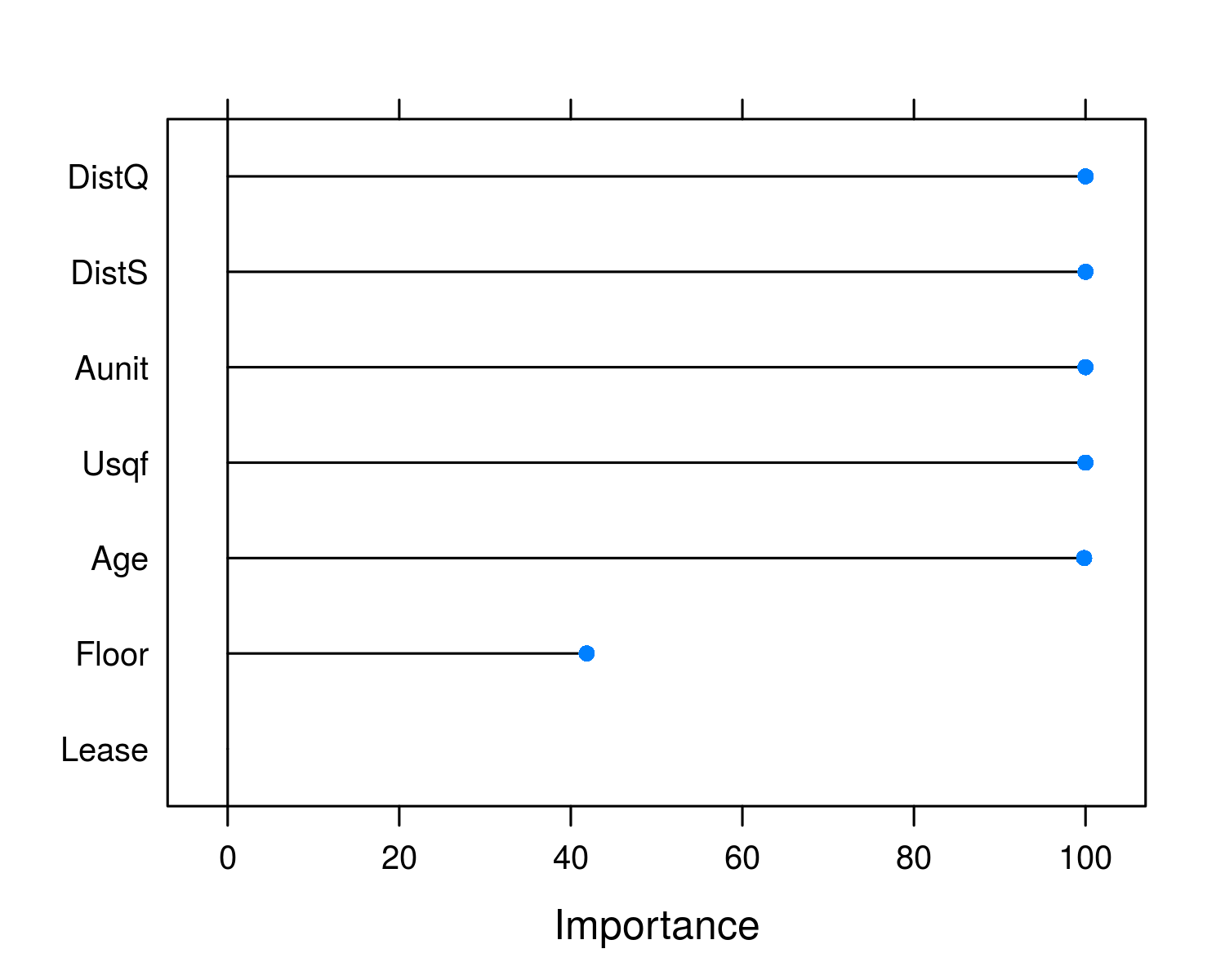


## Tuning Parameter Selection Using LOOCV



From above plot, lambda =100.00 and s =0.35 gives the minimum RMSE model.

## Variable Importance



## Standardized Model Coefficients

| **Variable** | **Estimate** |
| --- | --- |
| Age | -0.450 |
| Floor | 0.000 |
| Lease | 0.150 |
| DistQ | -0.368 |
| DistS | 0.338 |
| Usqf | 0.000 |
| Aunit | 0.135 |

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** |
| --- | --- | --- |
| 3.094 | 0.338 | 0.879 |