DSI - Project 3

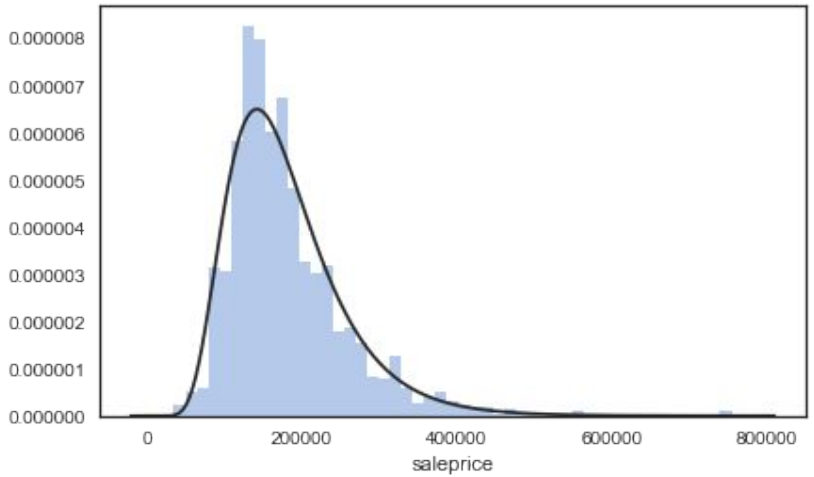
Ames Housing Data

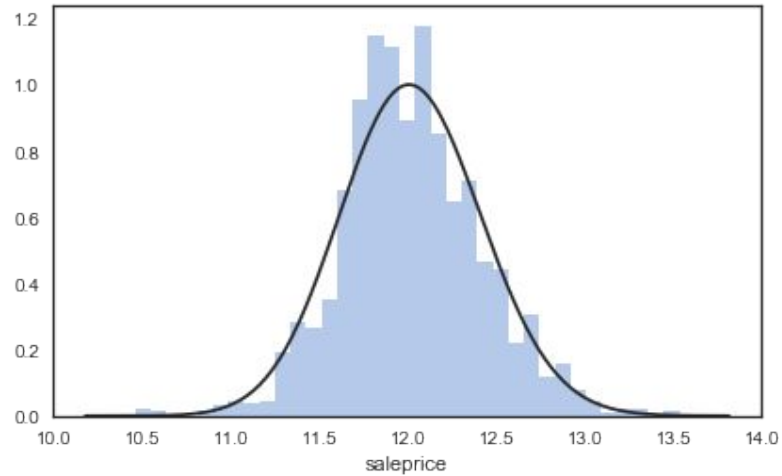
numpy.log1p

Calculates log(1 + x)

Original:
Kurtosis: 6.502799105301849
Skewness: 1.880007613681095

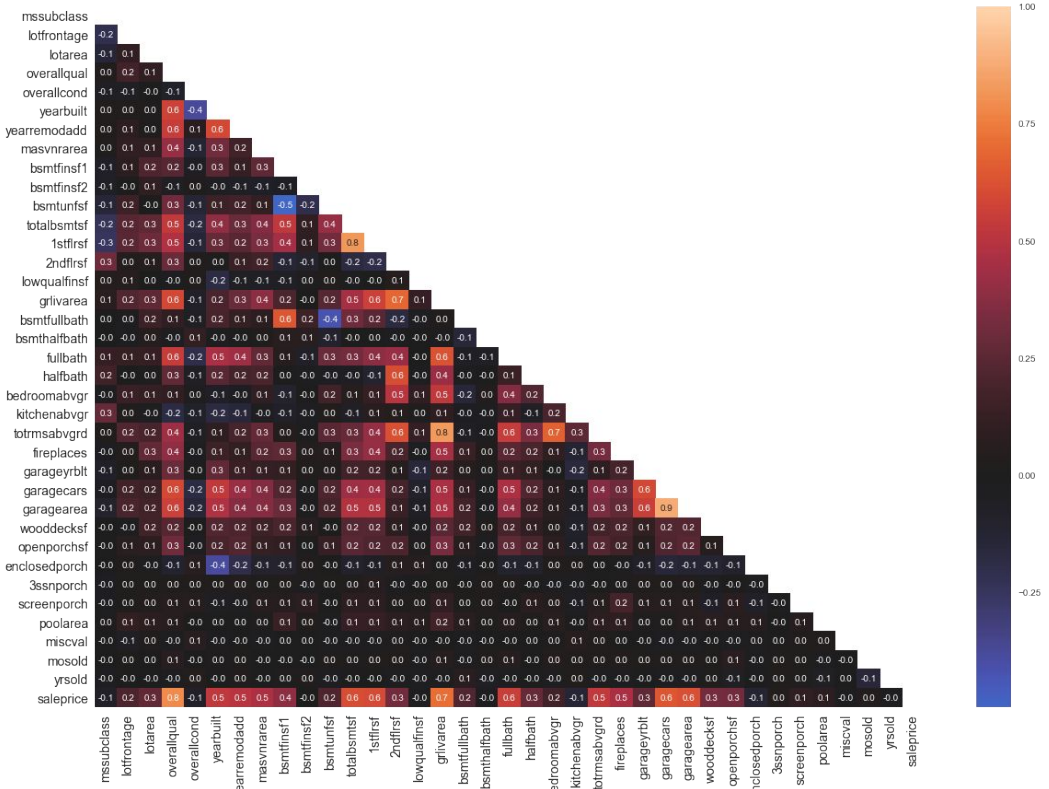
Fit:
Kurtosis: 0.8000427774841801
Skewness: 0.12120453554324111





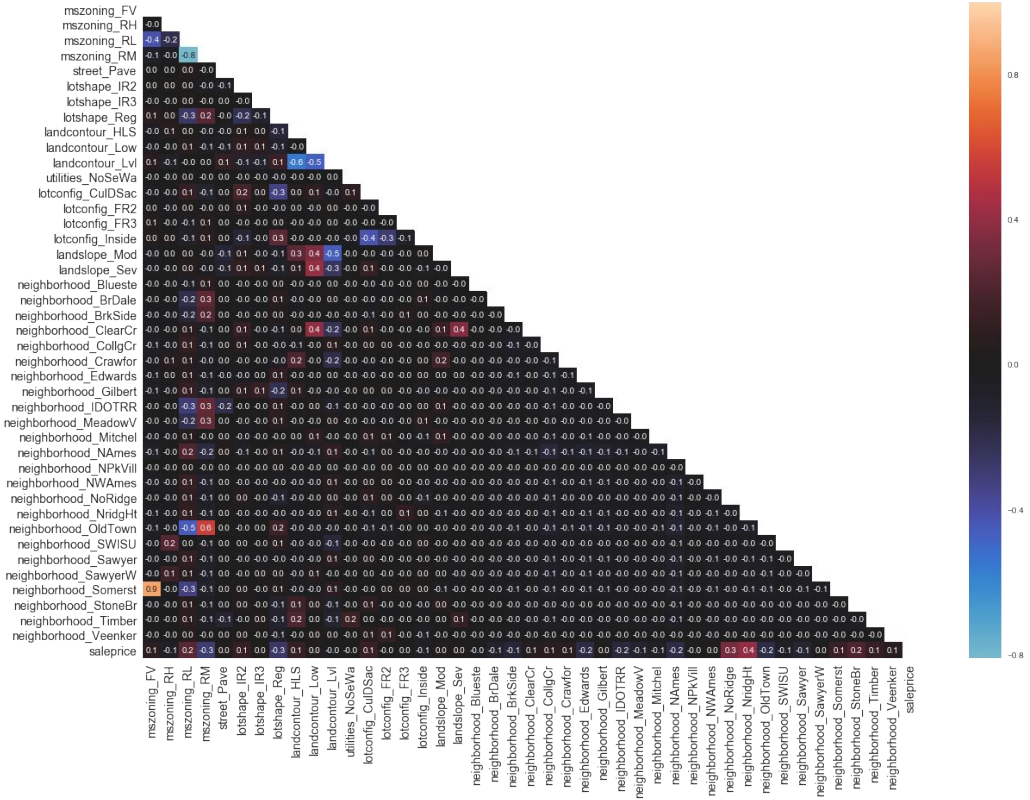
Quantitative heatmap

- Finding which quantitative features correlated with sale price the most.
- Seeing which features were heavily correlated with others.



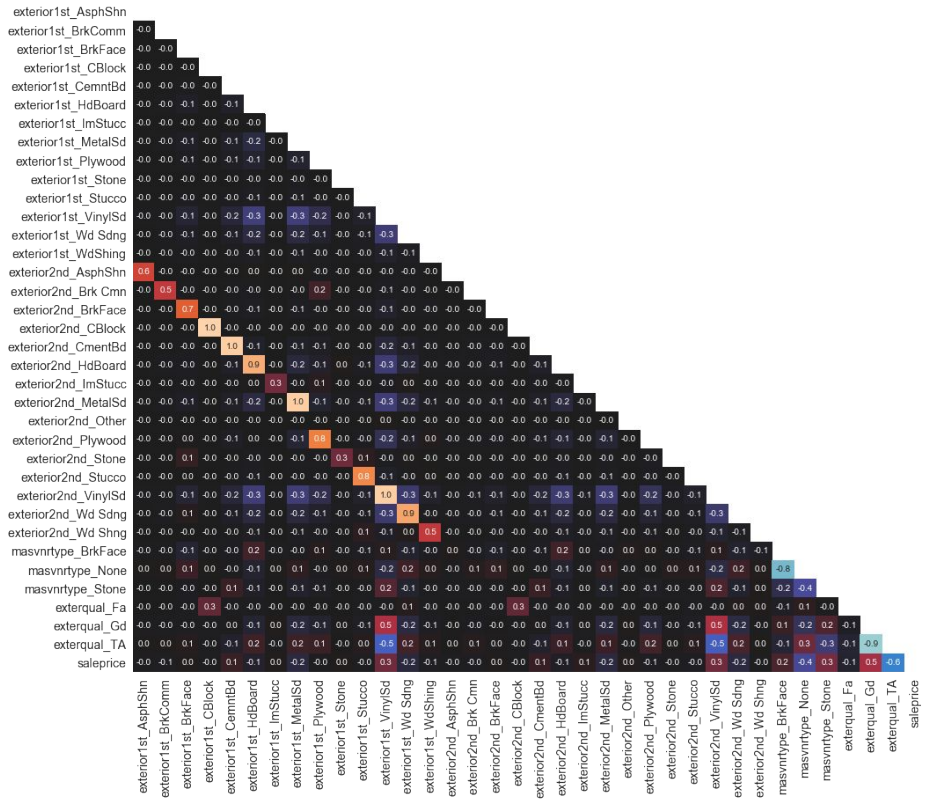
Comparing dummies

- Unique values for every qualitative column was counted to determine which features would be used in each heat map. This is because a heat map of all the dummy variables would be too large.
- Dummies were created from qualitative data.



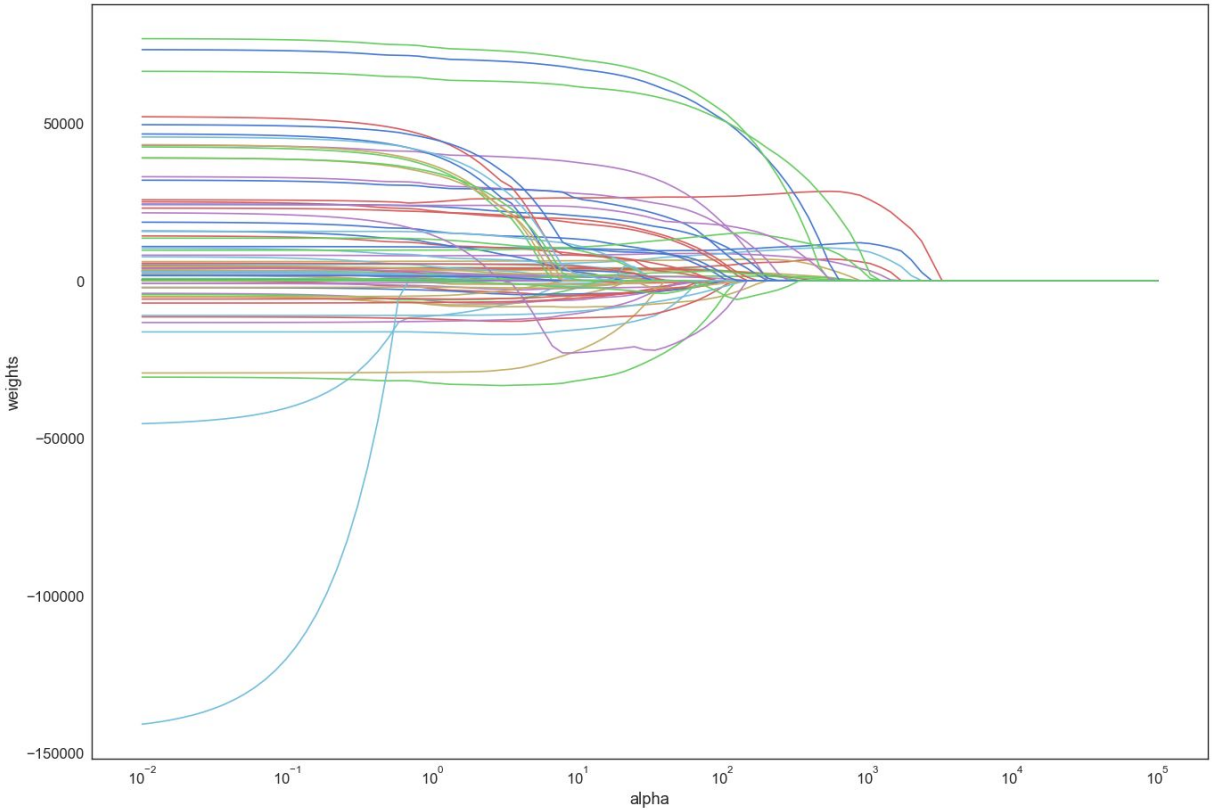
Comparing dummies

- Unique values for every qualitative column was counted to determine which features would be used in each heat map. This is because a heat map of all the dummy variables would be too large.
- Dummies were created from qualitative data.
- A few of the dummy heat maps showed that some categories were heavily correlated with others which was considered in the feature selection process.



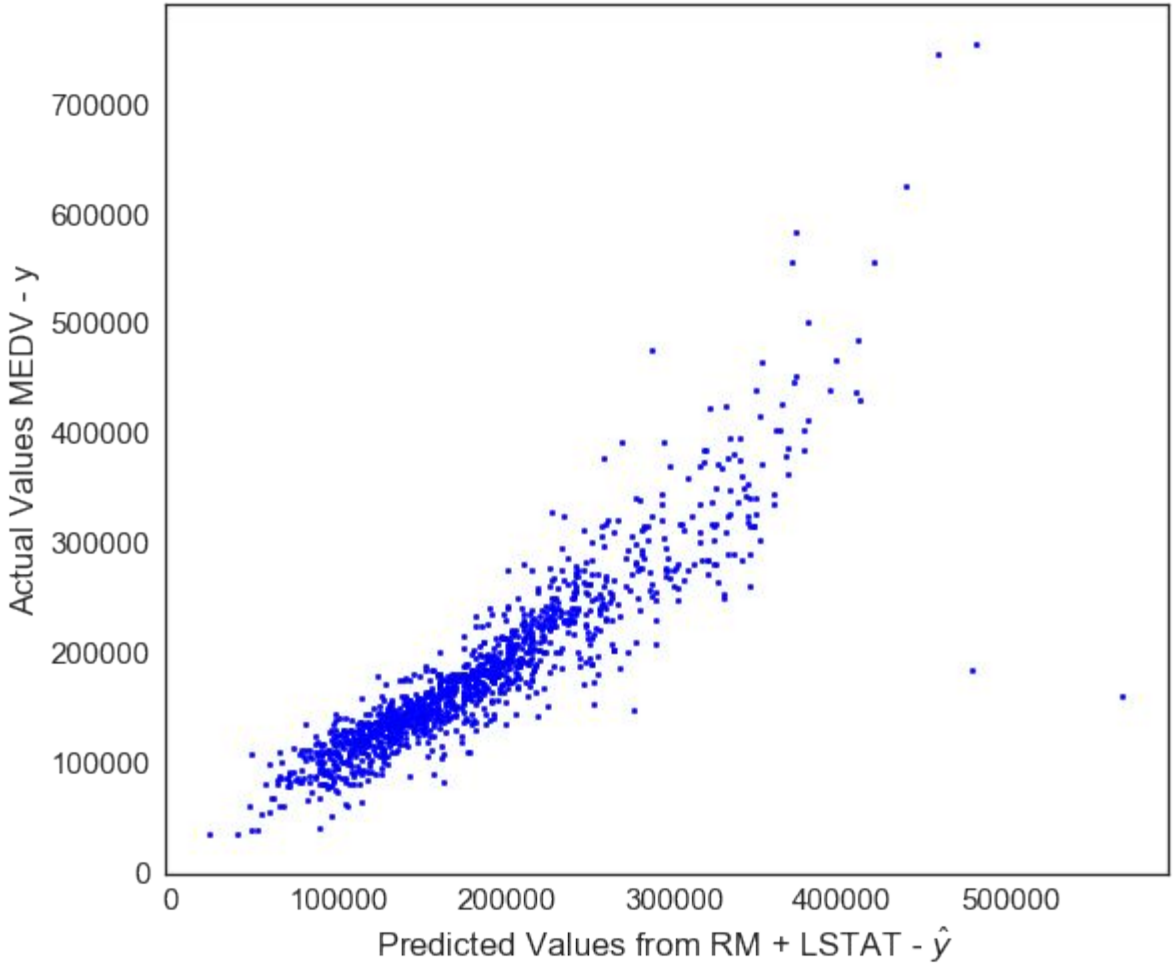
Lasso regression

- 59/81 columns were selected from Lasso regression.
- The names of these columns were then appended to a list and ordered by the highest coefficient value.



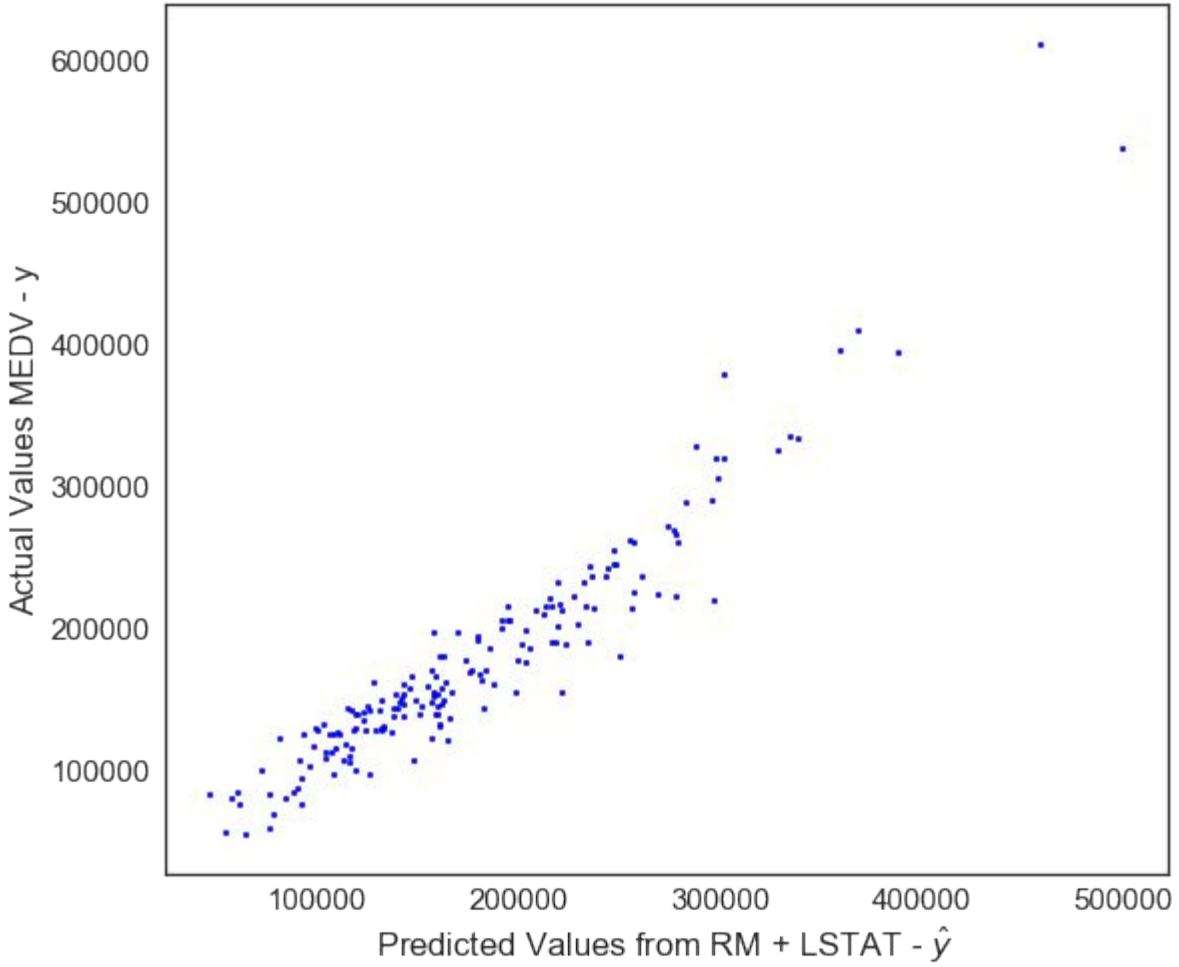
Lasso regression

Train scores:
R^2 of the prediction: 0.8191
Mean Square Error: 1138536014.24
Variance: 0.14747563830413407



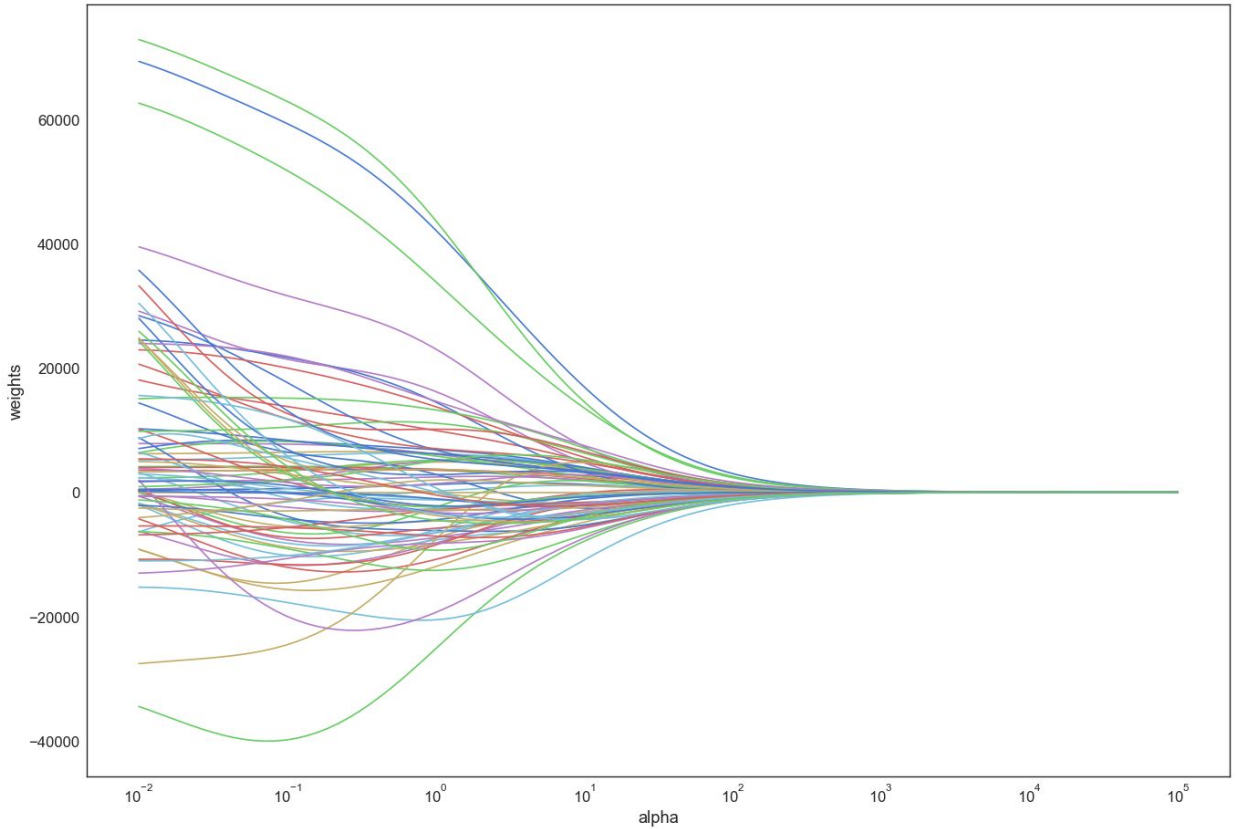
Lasso regression

Test scores:
R^2 of the prediction: 0.9040
Mean Square Error: 617997352.05
Variance: 0.10509698884293482



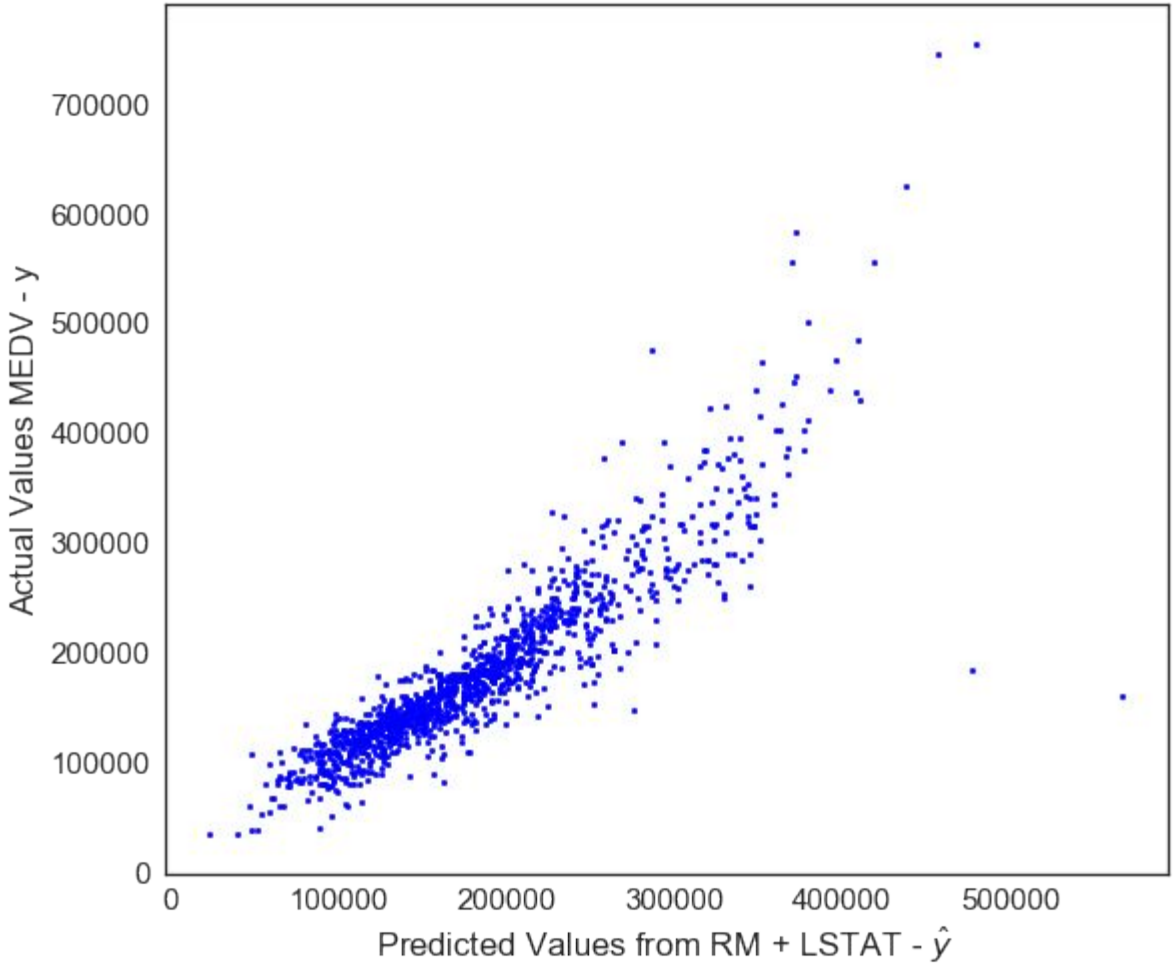
Ridge regression

- 54/81 columns were selected from Ridge regression.
- The names of these columns were then appended to a list and ordered by the highest coefficient value.



Ridge regression

Train scores:
R^2 of the prediction: 0.8198
Mean Square Error: 1134110342.56
Variance: 0.14747563830413407



Ridge regression

Test scores:
R^2 of the prediction: 0.8921
Mean Square Error: 694379572.11
Variance: 0.10509698884293482