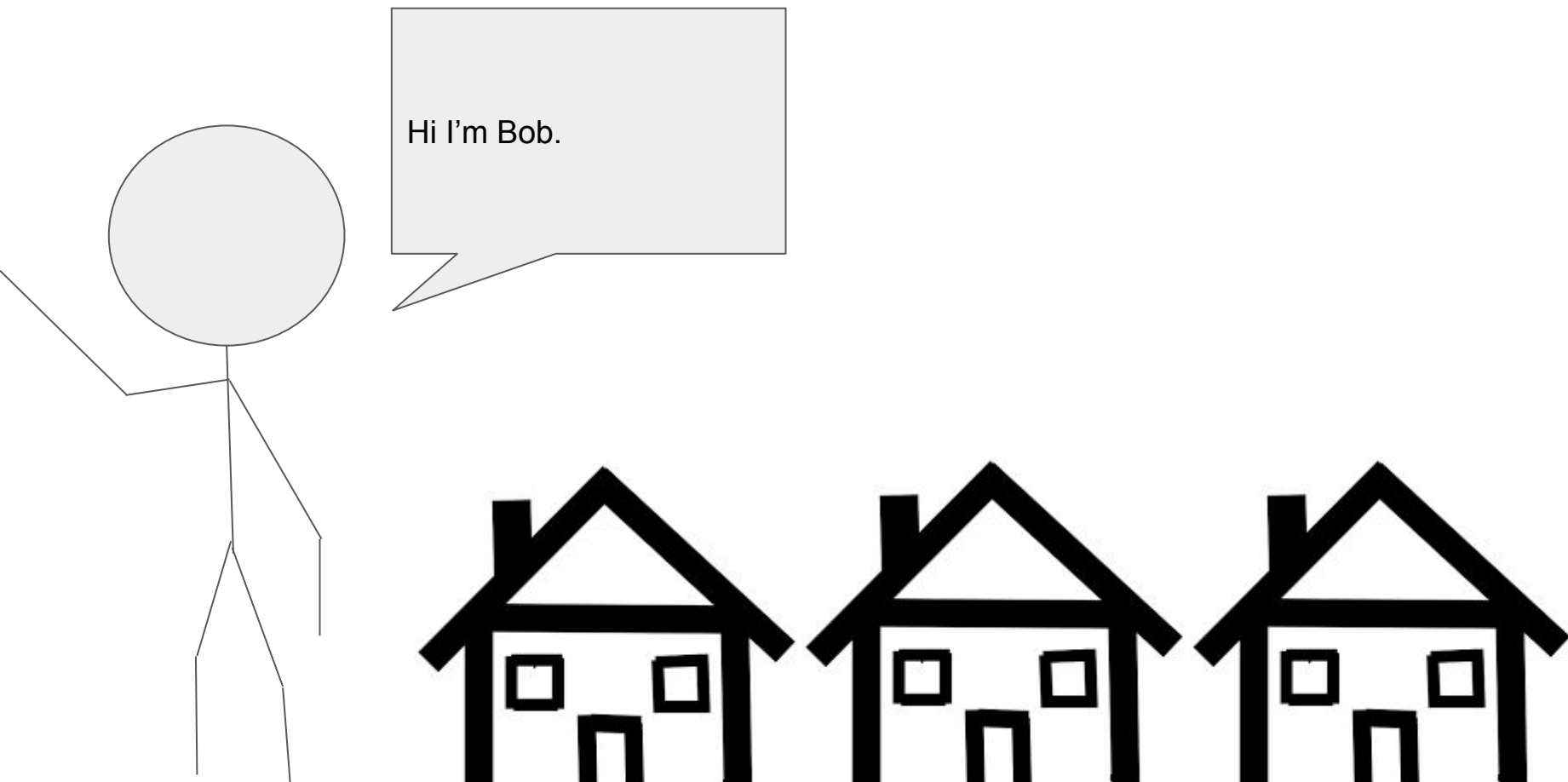


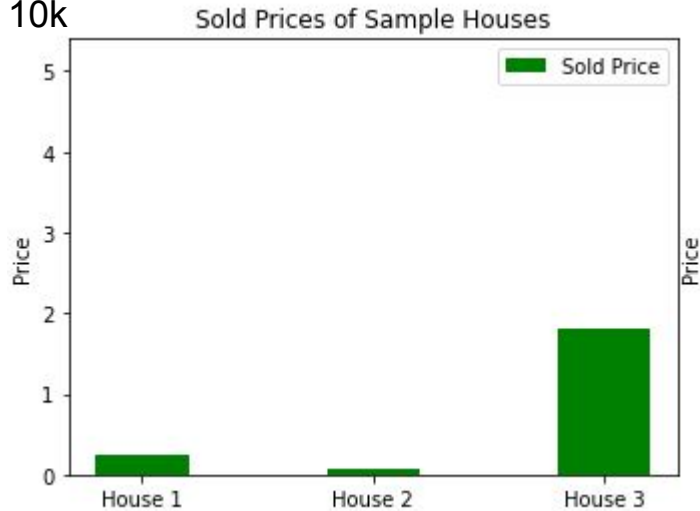
Linear Regression of House Prices before and after renovations

Elliot Ledson



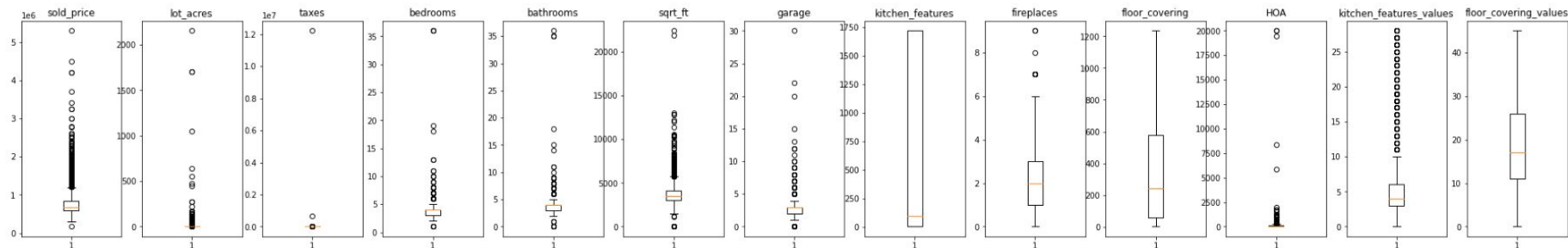
So below according to my model these 3 houses Bob acquired seemingly at random originally purchased at this price

Per
10k

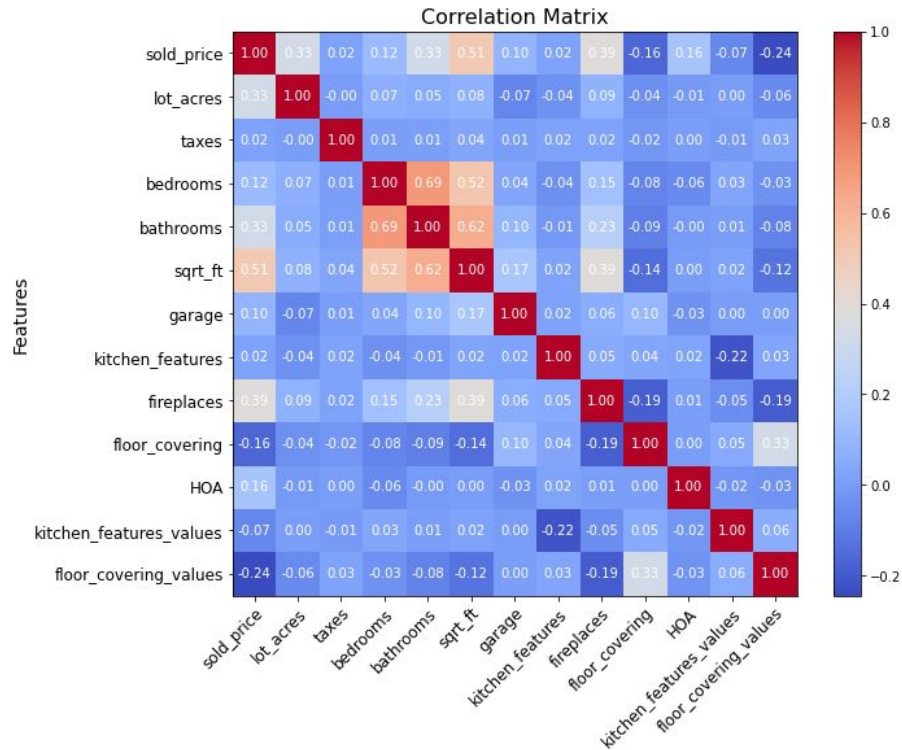


By the end of this presentation we will see a side by side comparison of what Bobs properties were purchased for and what they “could” be worth.

Bob's trusting me to find the best estimated projection for his properties after renovations. After some thorough preprocessing, I plotted box plots to check for outliers.



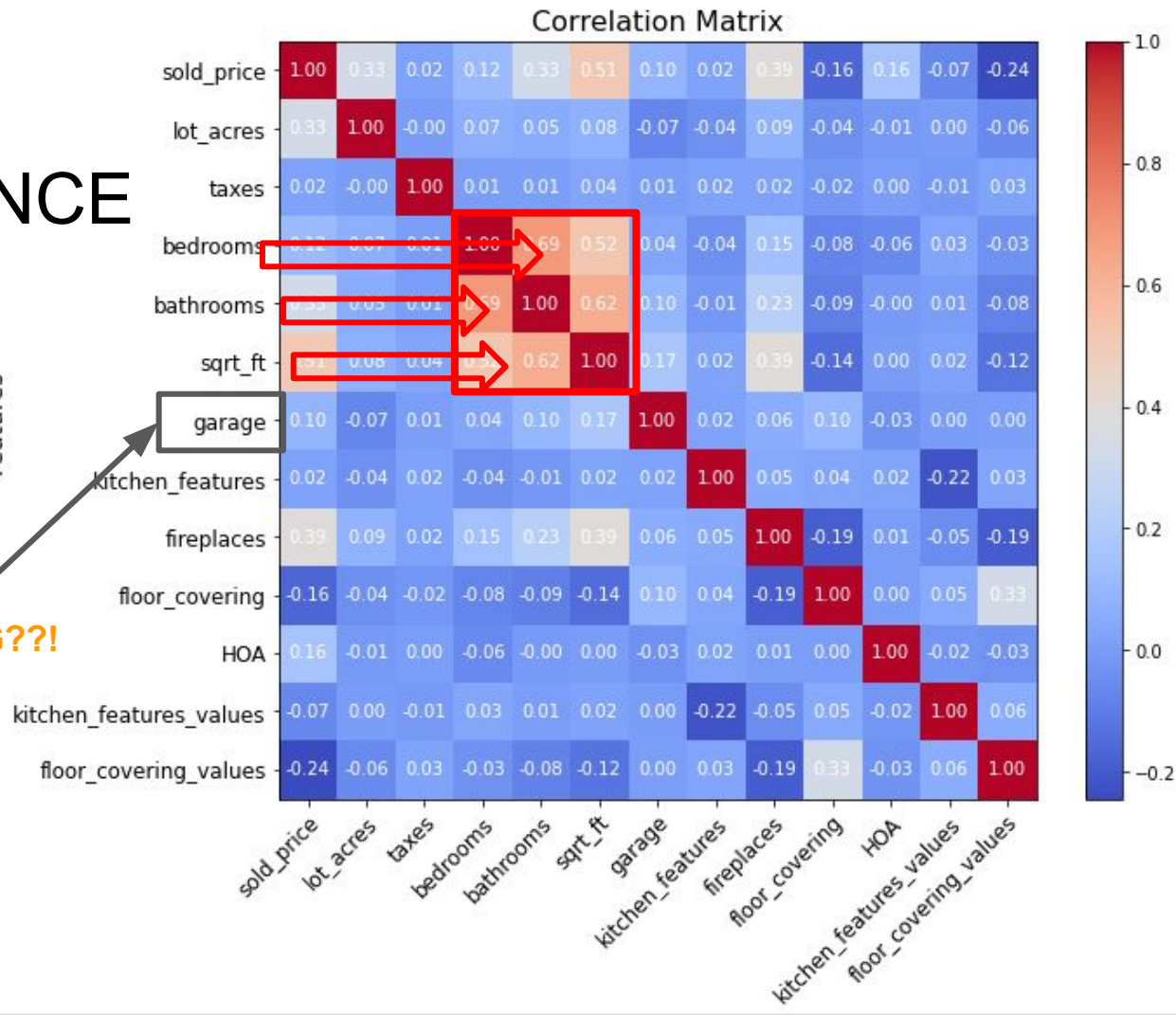
I implemented a correlation matrix



ENHANCE

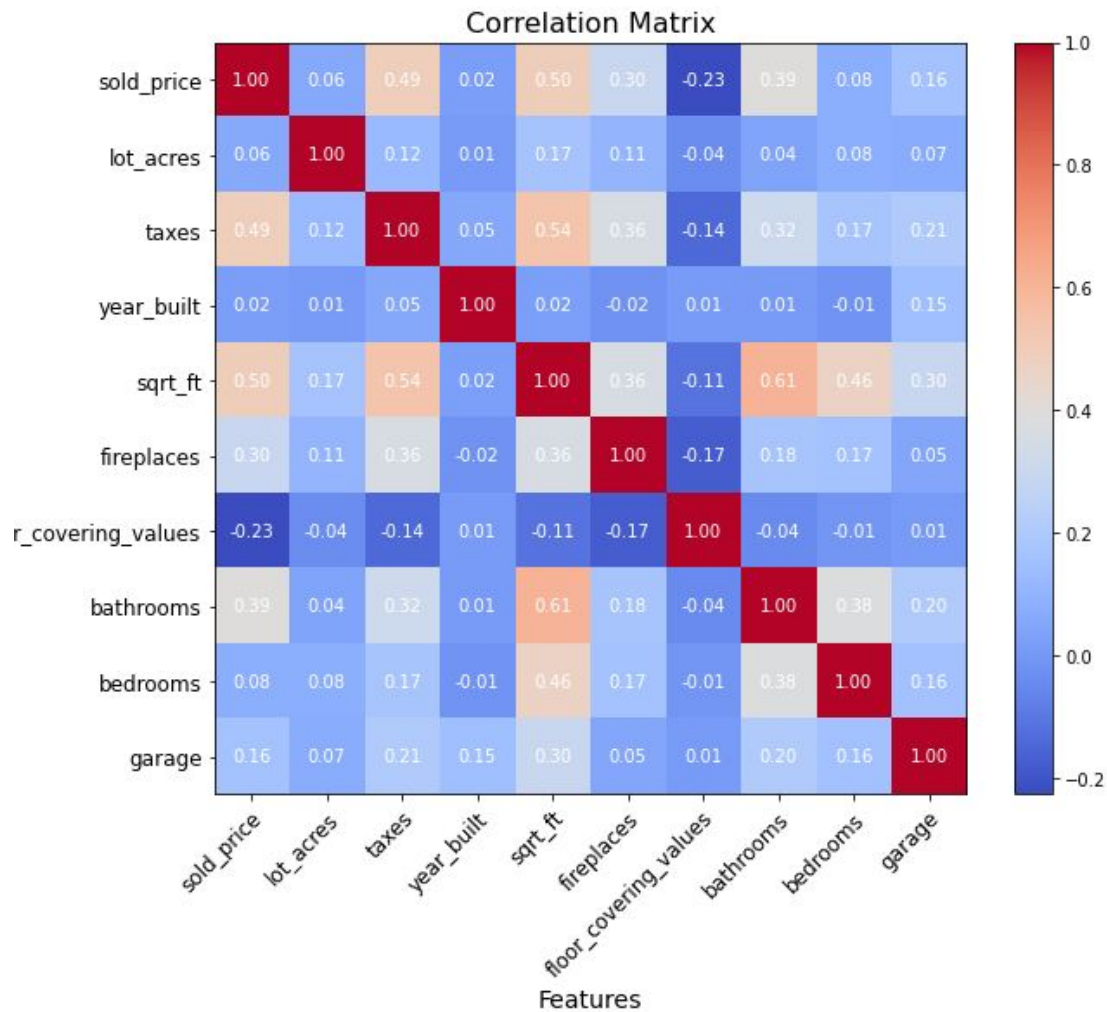
SURPRISING??!

Features



I ended up removing outliers that were greater than 3 standard deviations from the norm.

Going back I think I would probably tweak this part but at some point I have to accept what I have and learn from it.



My strong correlations disappeared

I was determined to get this done and it was like after midnight at this point.

I ended up using an 80:20 split for testing:training my data

I employed the multivariate linear regression algorithm

I adjusted the internal weights with the fit method

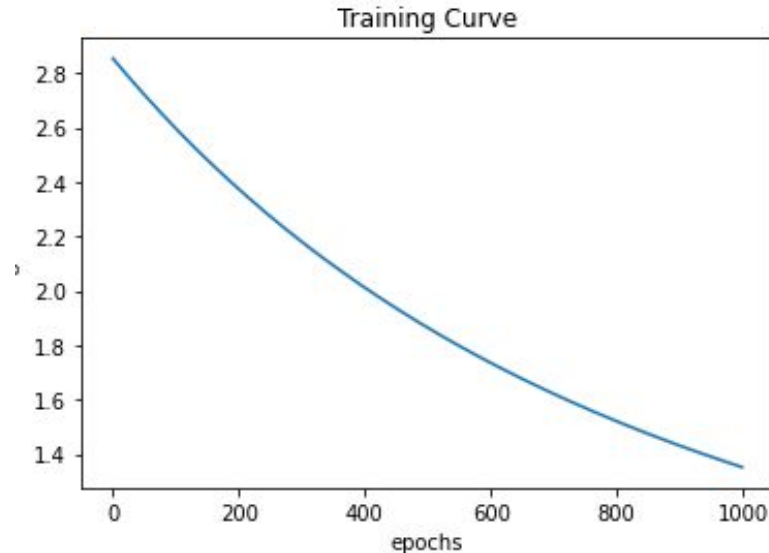
My predict method generates predictions

OLS() method is a helper function that calculates the error between the predicted and actual target values, and it's used by the fit() method to update the model weights.

Fitting the linear regression (training the robots)

The machine fits the line passing in the trained data points targets labels. It also sets the `show_curve` parameter to `True`, which means that a plot of the training curve will be displayed after the model is trained.

During training, the model learns to predict the target values from the given features by updating its internal weights. Once training is complete, the model will be able to predict the target values of new data that it has not seen before. The training process is iterative and aims to minimize the difference between the predicted and actual target values.



The difference

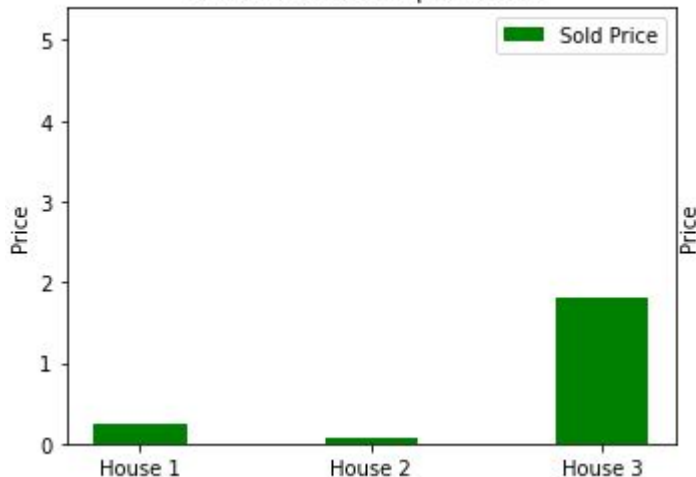
I am aware the x-axis counts are different I spent an hour trying to fix it.

Although it could use some refining, it is apparent that when taking in to account and adjusting the amount of renovations the house has.

'fireplaces', 'floor_covering_values', 'bathrooms', 'bedrooms', 'garage' that it makes a substantial difference to the price of the property

Per
10k

Sold Prices of Sample Houses



Price After Renovations of Sample Houses



I have just realised my model has an error in it that will need addressing.

I scaled extra values

```
'sqrt_ft', 'fireplaces', 'floor_covering_values', 'bathrooms', 'bedrooms', 'garage' 'taxes',  
'year_built', 'sqrt_ft'
```

Which has been factored in to my model, I would go back improve that when possible.