Don't *Flip* Out but...

Analysis by Michael Fedell, Dhansree Suraj & Tanya Tandon

Aim

To maximise profit assuming
Profit = Final_value - (Inital_price + Investment)

Executive Summary

Our group has been tasked with developing an analytical system for maximizing profit in our house-flipping business. Profit is captured in two ways: 1) identifying "underpriced" houses whose true value is greater than list price and 2) investing appropriately in houses so that final sell price is increased substantially. Our approach focuses on the former by modeling the true value of houses to identify advantageous opportunities.

Data Preparation

Fortunately, data is available on a per-house basis and no aggregations are necessary. However, of the 40 available predictors, the most relevant ones must be chosen and encoded appropriately. All numerical values were normalized by standard deviance from the mean, categorical predictors were treated as factors. Finally, only those predictors which were mostly complete (few NA values) were kept for the model.

Methodology

We approached this as a three tier problem. As the initial problem stated that two ways to successfully flip a house is to :

- Spot Undervalued Houses
- Upgrading the House

1. Initial Value Model

We postulated that the initial or “true” value of a property was likely a nonlinear function of the properties features and thus attempted various neural network architectures. However, these models did not prove to be significantly better than the baseline linear regression, and thus the multiple simple linear model was ultimately used with a high calculate R2 of 0.882.

2. Investment Model

The data yielded no significant relationships between reliable predictors and investment amounts made. However, we clustered investment purchases into several broad categories, and then used the average investment price for each category to determine the amount to invest on each new purchase.

3. Final Value Model

For predicting our final value, we saw that 216/218 cases in the training set were we sold at loss, the initial prices were more than the initial value of of the house. Therefore, we decided inital_value as the most important criteria. We filtered the selection of houses by inital_value > inital_price. After we filtered the house, we used only the inital_value calculated and investment calculated to decide on our final prices using a linear model and found our model had a R2 of > 0.75

Finally, we calculated the profit for each house based on the model we created and selected top k houses which satisfied our two constraints of budget.

Results

Our modeling techniques led us to allocate our $400,000,000 purchasing budget towards the acquisition of 1209 new properties. We spread our $100,000,000 investment budget among these properties to bring final sell price revenues up to $635,545,566.80.
Total Profit: $181,172,638.10