

Real Estate Rental Selection Augmentation  
Algorithm



# PricePoint

(High Value Markets Version)

# High Value Markets

- Definition: Densely populated markets where land and sq ft is at a premium
- Residential Price/Sq Ft Rates
  - New York - \$1397
  - San Mateo - \$757
  - Washington, D.C. - \$477
  - Arlington, VA - \$455

# Overview

- Question: In high value markets which property to buy and renovate for short term gain and the influencing attributes that generate greatest profit
- General hypothesis: Take a small ugly box and make it into the biggest prettiest box possible (given zoning considerations)

# Proposed Scenario

- 50 year old townhome never updated, 20ft x 30ft x 2 stories livable space (1200 sq ft)
- Newly renovated home 20ft x 50ft x 4 stories livable space (4000 sq ft)
- Arlington
  - Old Home - \$546K
  - Construction Costs - \$1M
  - New Home - \$1.8M
  - Profit - \$254K





# Additional Considerations

- Zoning Laws
- Sub Markets (Georgetown v. Columbia Heights v. Anacostia)
- Lot Size
- Current Condition
- Layout/Dimensions

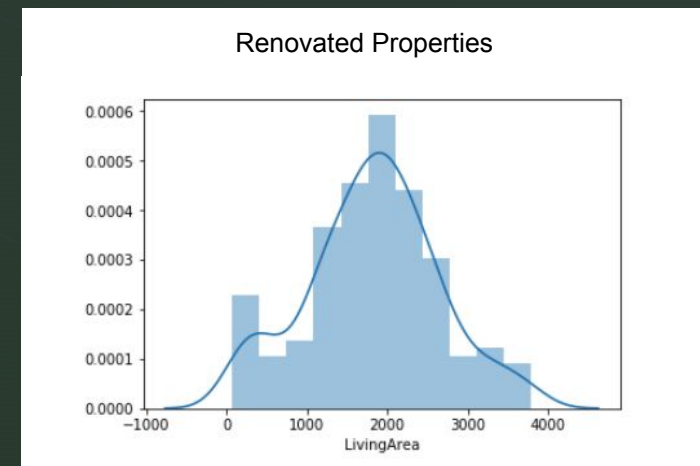
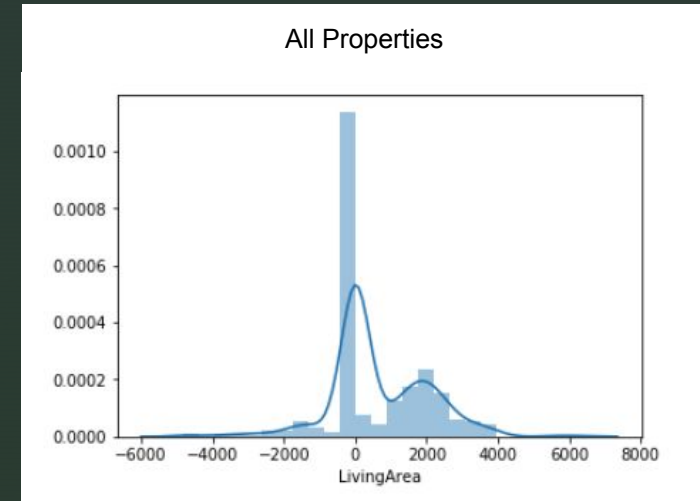
# Analysis Overview

- Attributes
  - All variables are deltas between last and most recent sales of a property
  - Independent variables - 53 features describing each property
  - Dependent variable - revenue gain for each property
- Modeling
  - Linear regression
  - Multi linear regression
  - Decision tree



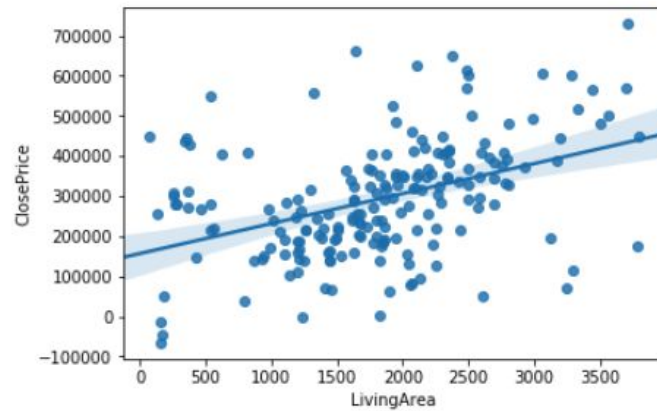
# Exploratory Data Analysis

- What typically changed in these houses being renovated
- Attributes that are traditional predictors of home pricing analyzed to see how changes in those attributes effected change in sales price

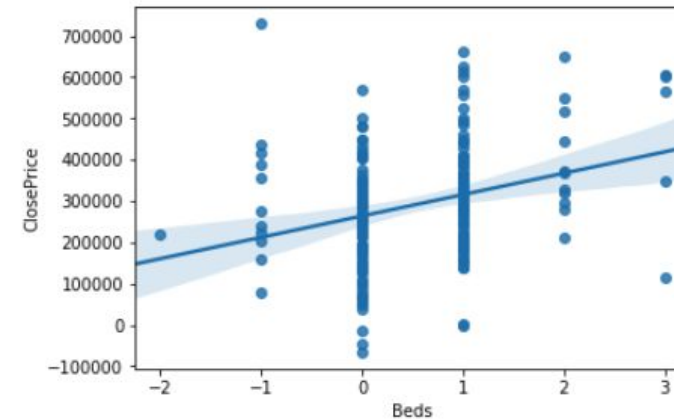


# Exploratory Data Analysis

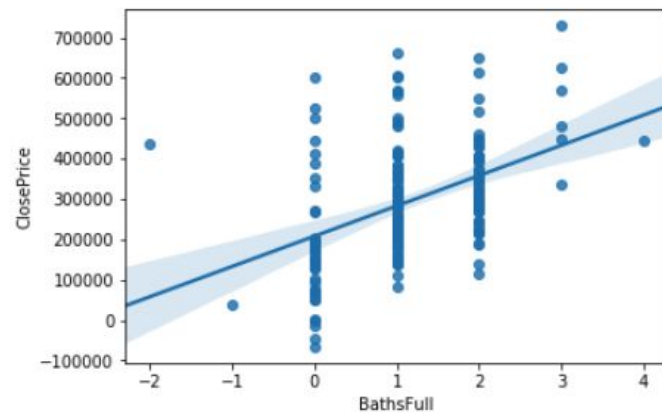
Change in Living Area to Change in Close Price



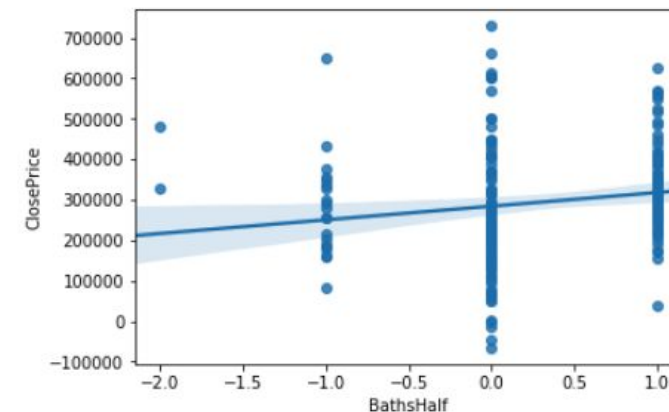
Change in Beds to Change in Close Price



Change in Full Baths to Change in Close Price



Change in Half Baths to Change in Close Price





# Multi Linear Regression

- Changes in Sq Ft, Beds, and Baths to ClosePrice

- Full Bath = \$72K
- Half Bath = \$42K
- Bed = \$25K
- Sq Ft = \$94

## OLS Regression Results

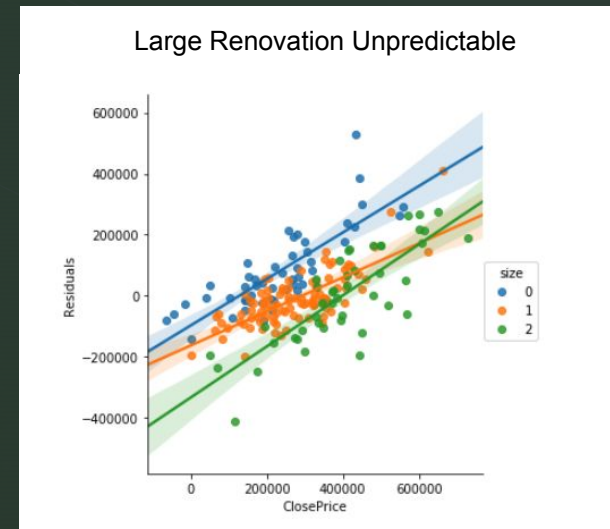
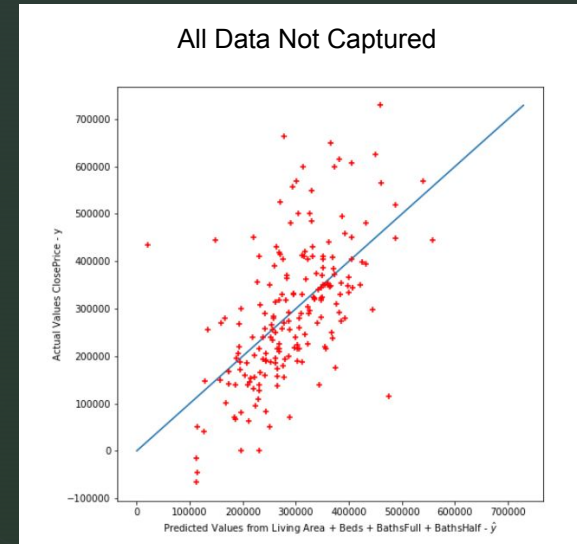
<b>Dep. Variable:</b>	ClosePrice	<b>R-squared:</b>	0.857
<b>Model:</b>	OLS	<b>Adj. R-squared:</b>	0.854
<b>Method:</b>	Least Squares	<b>F-statistic:</b>	283.8
<b>Date:</b>	Tue, 10 Apr 2018	<b>Prob (F-statistic):</b>	5.98e-79
<b>Time:</b>	06:31:43	<b>Log-Likelihood:</b>	-2550.0
<b>No. Observations:</b>	194	<b>AIC:</b>	5108.
<b>Df Residuals:</b>	190	<b>BIC:</b>	5121.
<b>Df Model:</b>	4		
<b>Covariance Type:</b>	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
<b>LivingArea</b>	94.1857	7.573	12.437	0.000	79.248	109.123
<b>Beds</b>	2.527e+04	1.16e+04	2.176	0.031	2360.930	4.82e+04
<b>BathsFull</b>	7.153e+04	1.06e+04	6.726	0.000	5.05e+04	9.25e+04
<b>BathsHalf</b>	4.299e+04	1.31e+04	3.282	0.001	1.72e+04	6.88e+04

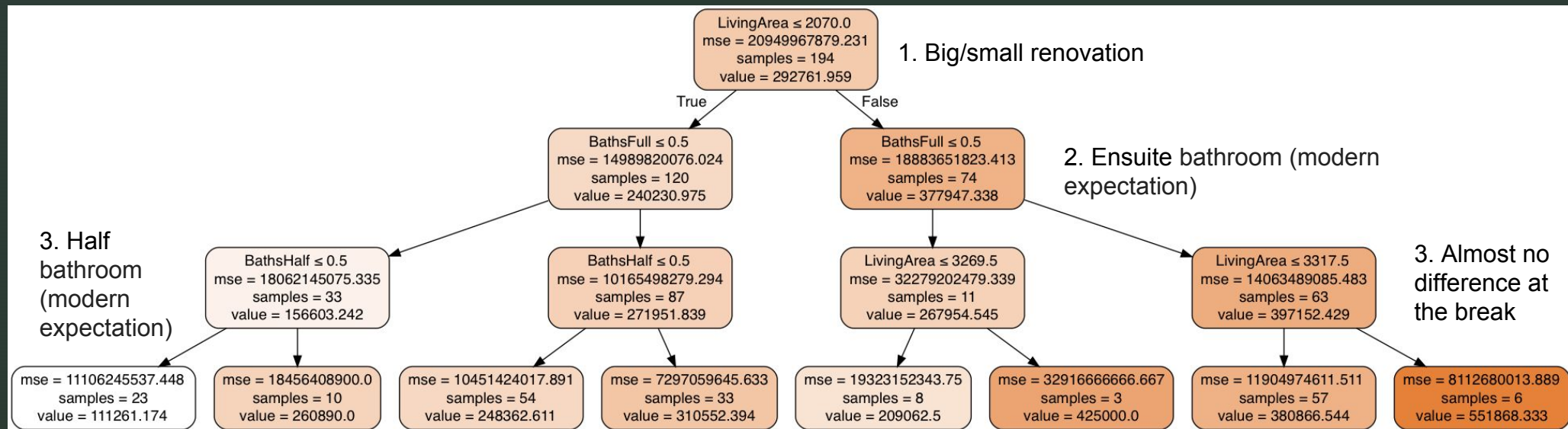
<b>Omnibus:</b>	28.132	<b>Durbin-Watson:</b>	2.133
<b>Prob(Omnibus):</b>	0.000	<b>Jarque-Bera (JB):</b>	53.770
<b>Skew:</b>	0.715	<b>Prob(JB):</b>	2.11e-12
<b>Kurtosis:</b>	5.147	<b>Cond. No.</b>	2.92e+03

# Multi Linear Regression

- Current model shows correlation but does not capture all the data
- Large renovations are more unpredictable



# Decision Tree



# Conclusion and Way Ahead

- Conclusion: Take a small ugly box and make it into the biggest prettiest box possible (given zoning considerations) and meet modern expectations
- Way Ahead:
  - Analyze additional attributes
  - Get a larger sample size via condo conversions
  - Get zoning data (what is in the realm of the possible)
  - Pull in cost data associated with different renovations
  - Geospatial analysis



# Thank You

