Project 2 - Ames Housing Data

DSIF3 Group 2

Welcome!

Property Flipping Masterclass

11 Dec 2011

Team Proprata

Shu

Cheng Yeow

Sean

Kris

Prop Flippin’

be

Flippin’ EZ

Today's masterclass

01

Introduction

Problem Statement

02

Our Secret to Success

Data Cleaning, EDA, Visualizations

03

Our Winning Formula

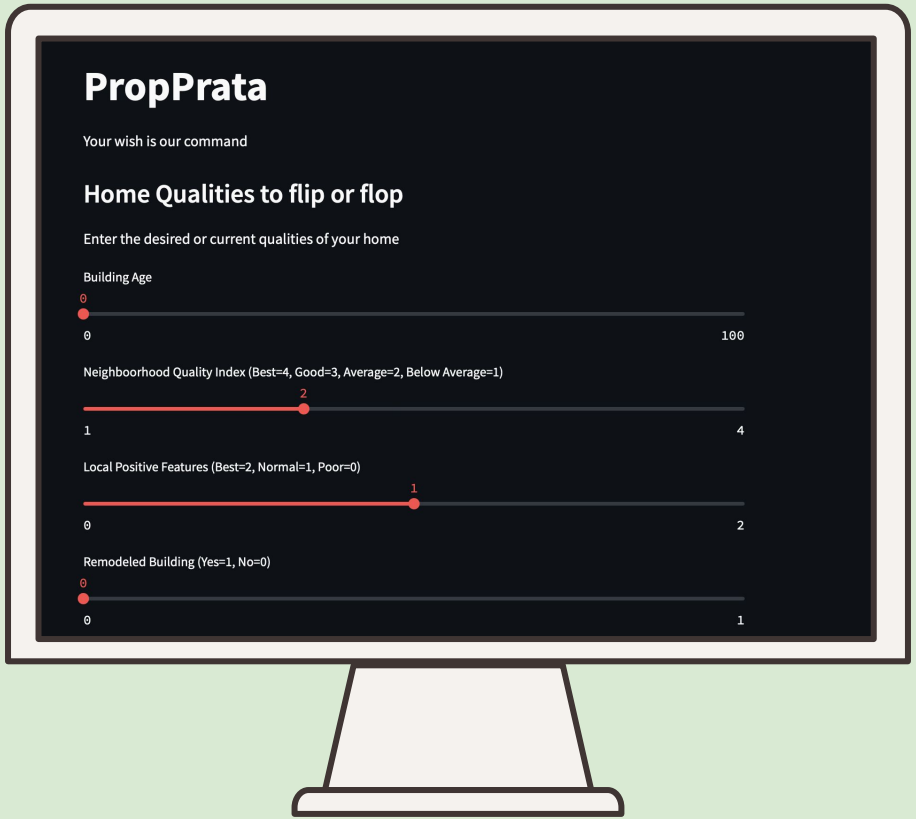
Modeling and Evaluation

04

What We Have to Offer

Conclusion and Recommendations

Live Software Demo



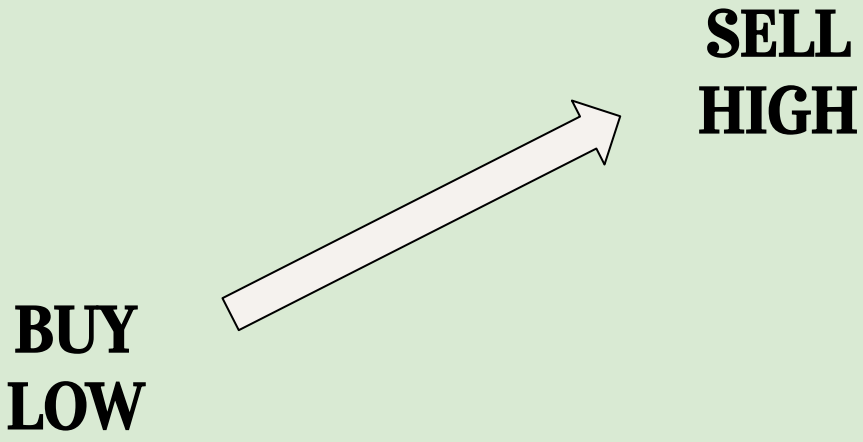
01

Problem Statement

Introduction

Introduction

Property flipping is an **investment strategy** of purchasing a property with a short holding period with the intent of selling it for a quick profit. ***All investments have risks***.



Two types of property flipping

Hot Market Conditions

1. Rapidly appreciating market
2. Little to no renovation
3. Higher risk

Renovation Flip

1. Buys undervalued property
2. Improves with renovation
3. Lower risk

Problem Statement

To help novice property flippers manage risks, we will provide consultation upon application, based on **our home valuation** that predicts the sale prices of homes in Ames, Iowa. We have modeled **Linear Regression, Ridge, Lasso and Elastic Net models**. The **sale price prediction software** is also helpful for property buyers or sellers.

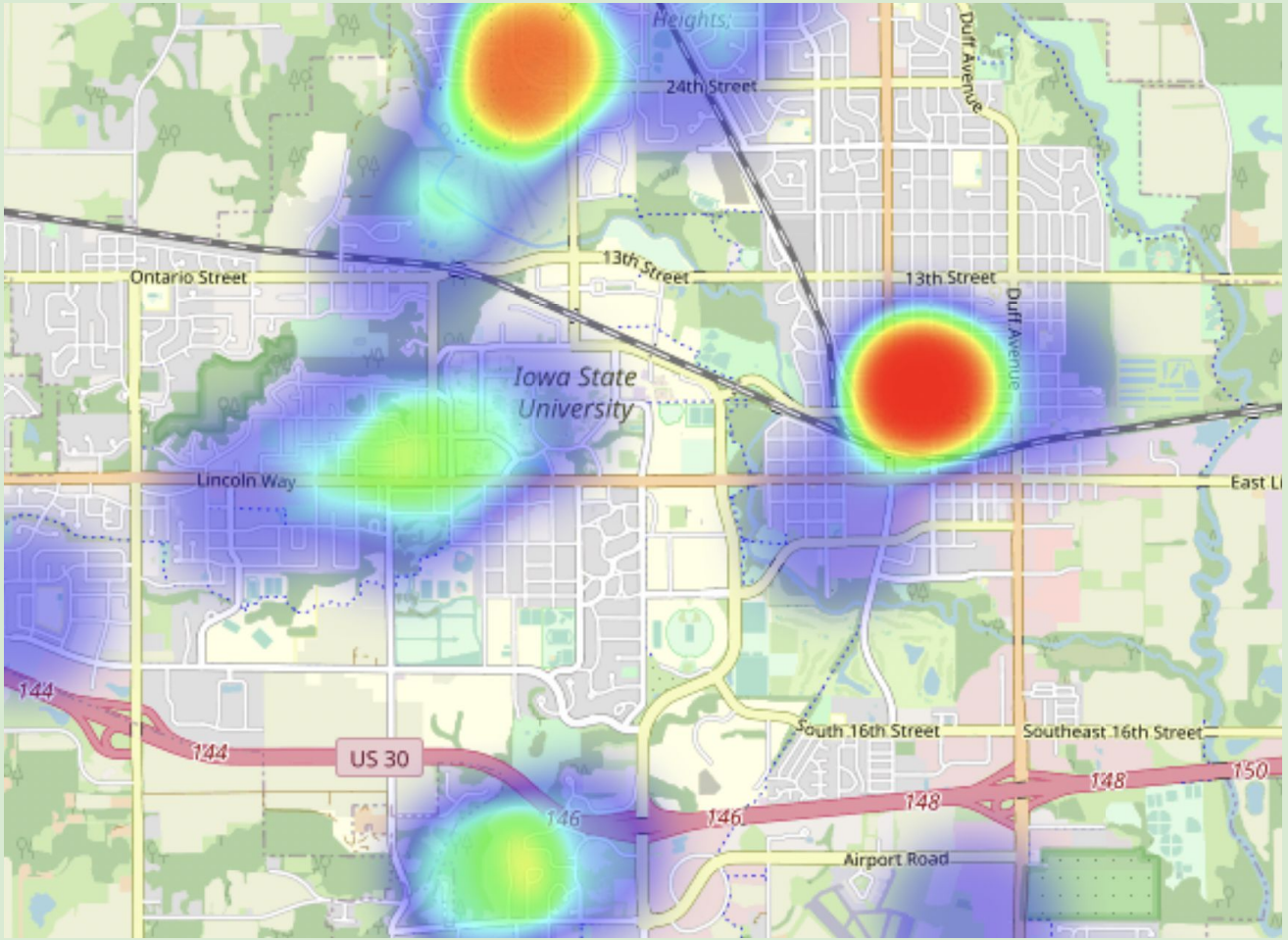
To **better manage risks** for new property flippers, we recommend the Reno Flip.

Thus, our recommendations will **focus on property features** that are more likely to **increase property value** when renovated, and which features to de-prioritise.

“Success in real estate starts when you believe you are worthy of it.”

Michael Ferrara

Location, Location, Location.



Heatmap of sale prices of properties in Ames.

02

Data Cleaning, EDA, Visualizations

Our Secret to Success

Exploring the data

2,051 * 80 = 164,080

Properties

Sale of Ames properties in 2006 - 2010

Features

Such as Basement conditions, kitchen quality.

Data points

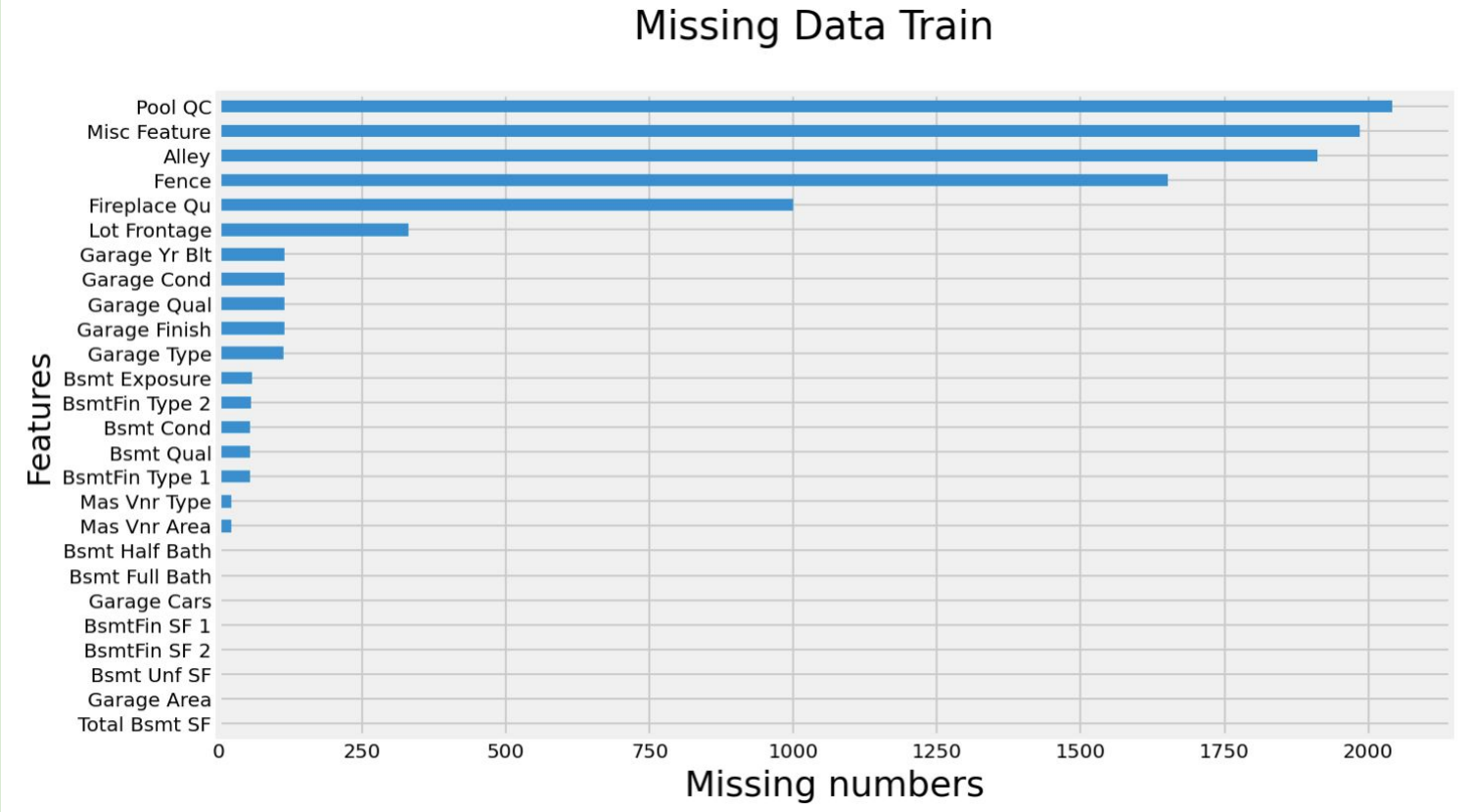
Bits of information

26

Features with **missing data**

9,822

Total **missing data**

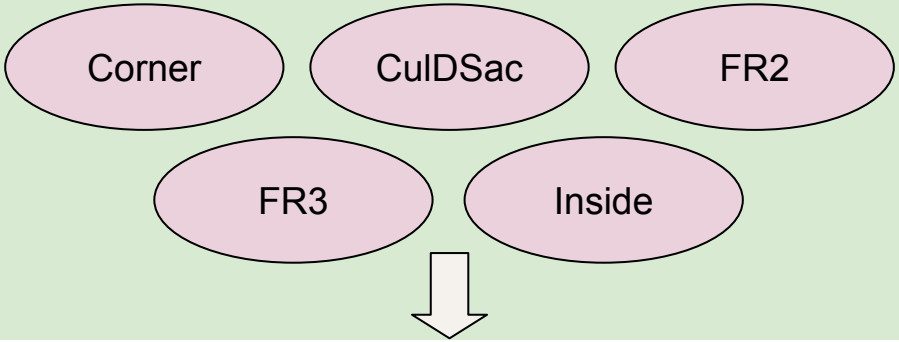


Dealing with missing data - Example

Lot Frontage

Filled missing values with median

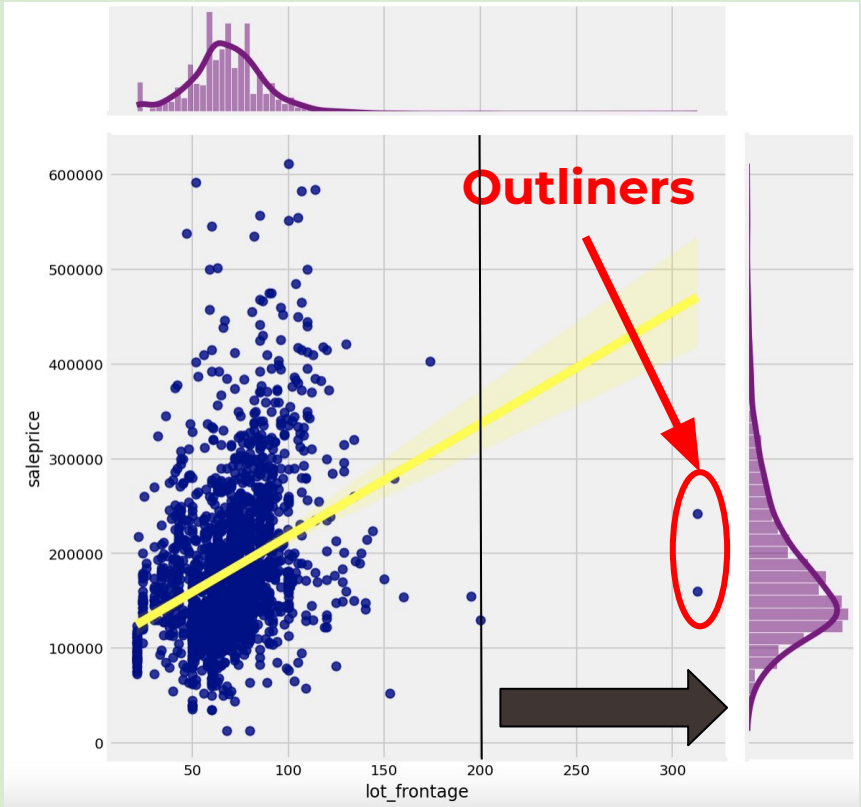
Neighborhood - Lot Config

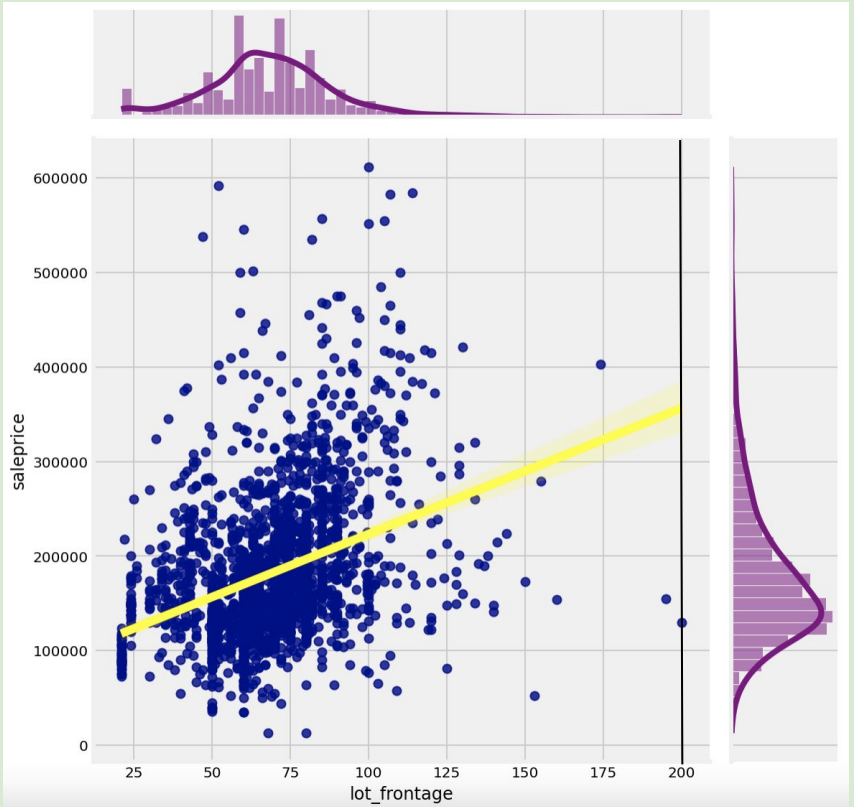


Transformation

Neighborhood	Blmngtn	Blueste	BrDale	BrkSide			
Lot Config	Inside	Inside	Inside	Corner	FR2	FR3	Inside
Lot Frontage	43.0	24.0	21.0	51.0	60.0	79.5	51.0
Lot Area	3189.0	1866.0	1680.0	6180.0	6911.0	6047.5	6240.0

Dealing with outliers - Lot frontage





Feature Engineering - Neighborhood

Type of variable	Transformation	Features
Ordinal	Mapping of text to numeric 'Ex': 5,'Gd': 4,'TA': 3, 'Fa':2,'Po':1	Neighborhood, Kitchen quality

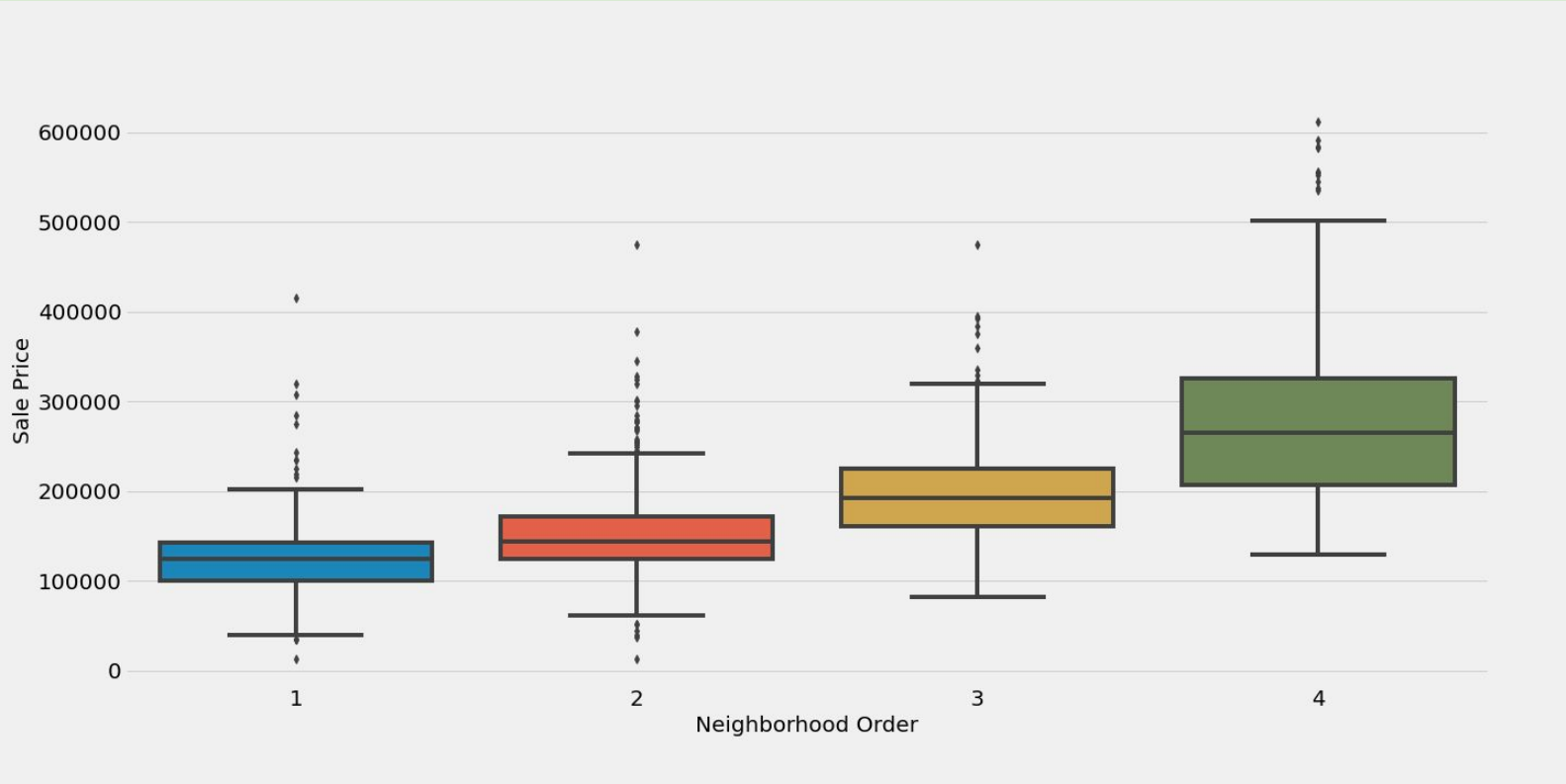
Feature Engineering - Example -KIV

Neighborhood Feature

Classify using the mean value of overall quality & condition, external quality & condition and functional features

Neighborhood Order	Avg Sale Price	No. of Properties
1	124,518	399
2	150,775	791
3	198,180	455
4	278,513	406

Feature Engineering - Neighborhood Example



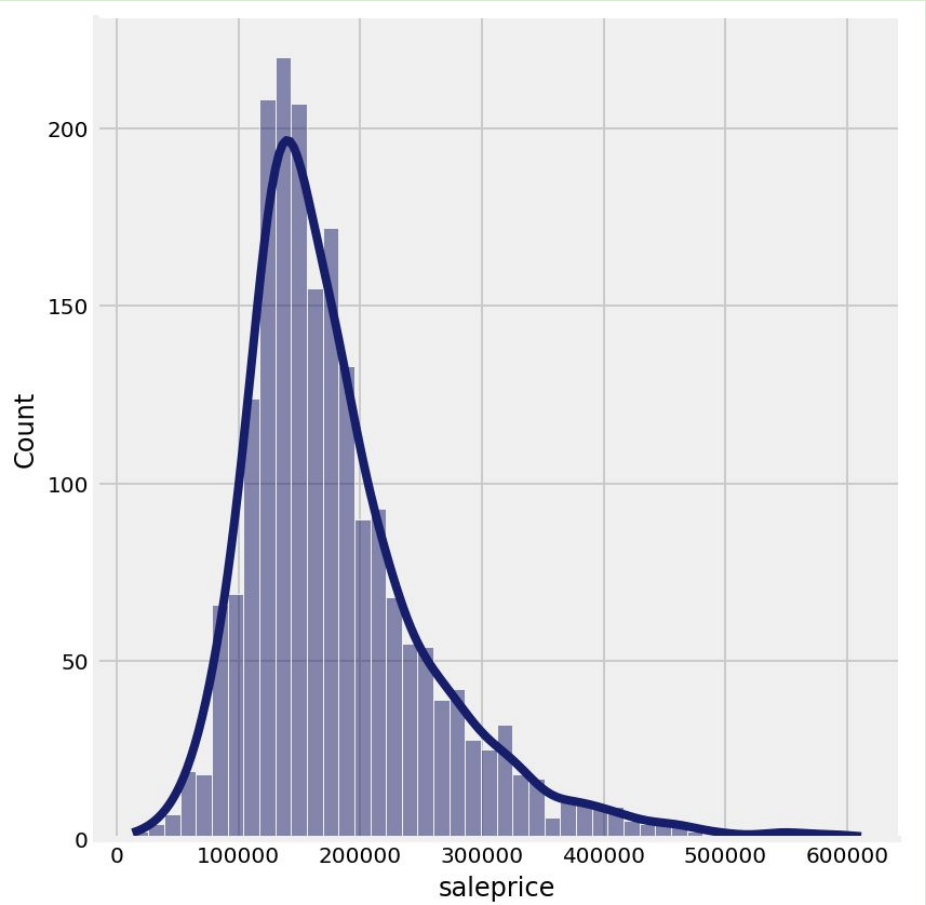
Cleaned Data

80

Original features

33

Useful features/
group of features



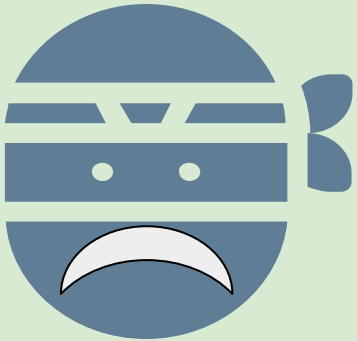
03

Modeling and Evaluation

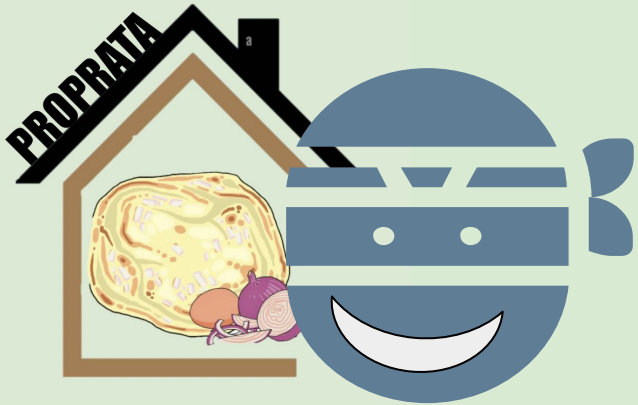
Our winning formula

Firstly.. A look at profits without ProPrata's help!

Losses in potential profits of up to $*70000 and more.*



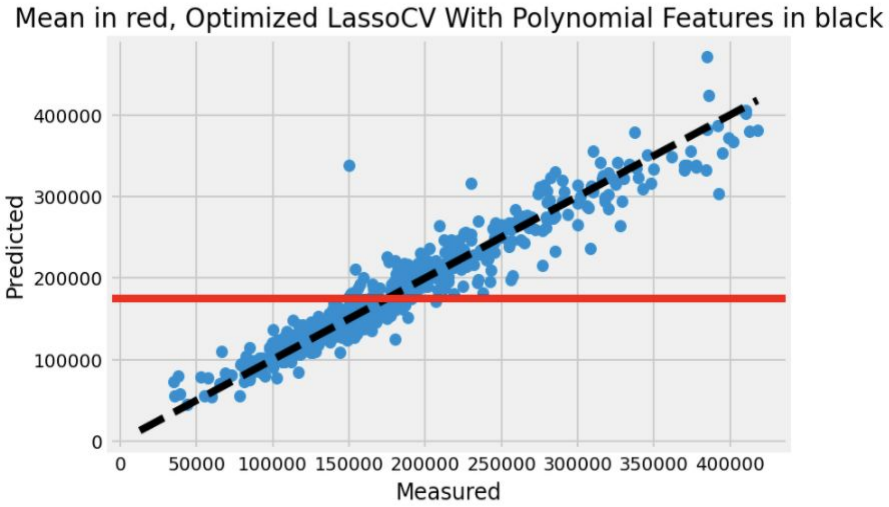
With ProPrata's help, profits have gone up as much as $48000!



But how did we do it?

Let’s take a look at our Proprata’s formula VS other companies’ formula!

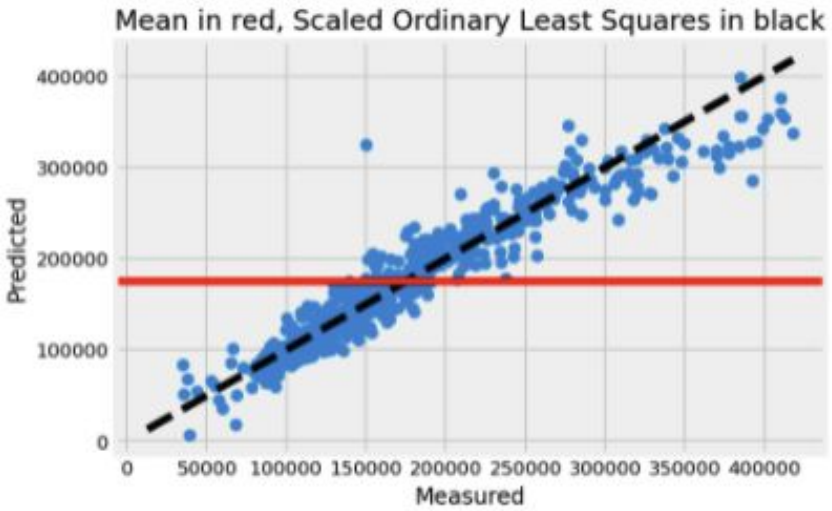
Proprata's Model



Prediction buffer
$20,253.00

Predictive power
past *clients* 92%
prospective *clients* 93%

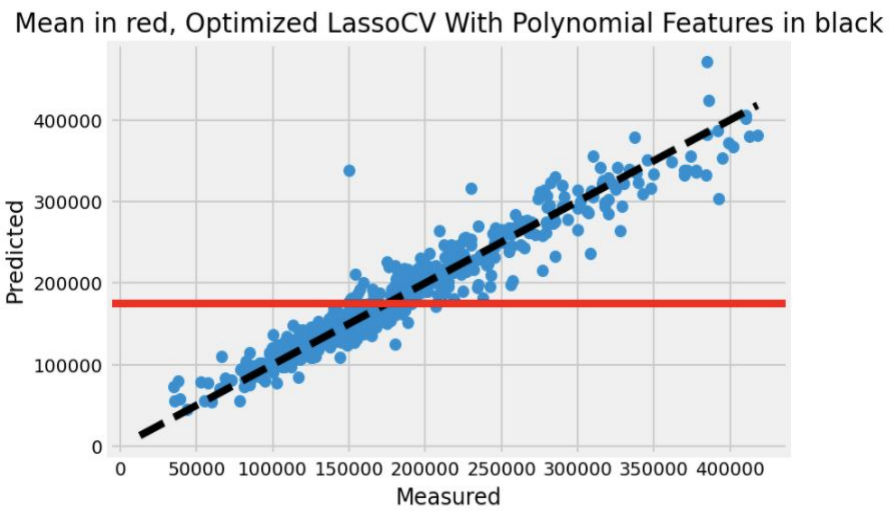
Coy A's Model



Prediction buffer
$22,692.69

Predictive power
past *clients* 89%
prospective *clients* 90%

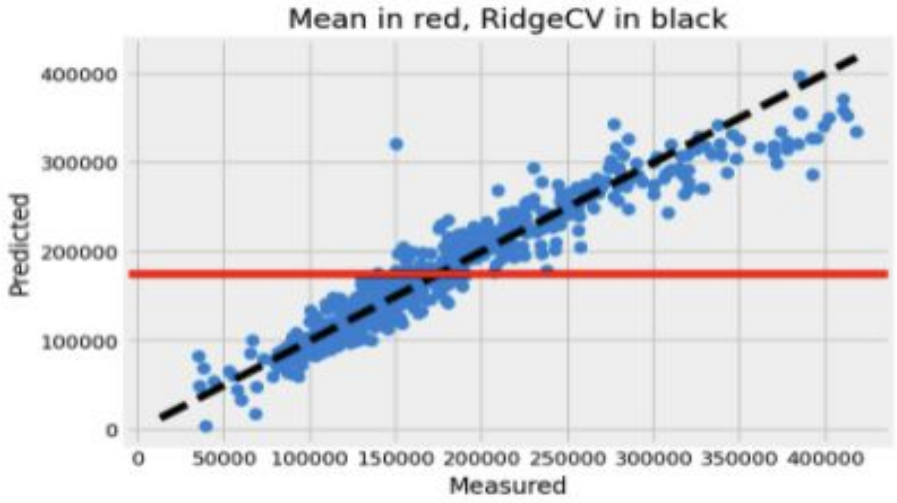
Proprata's Model



Prediction buffer
$20,253.00

Predictive power
past *clients* 92%
prospective *clients* 93%

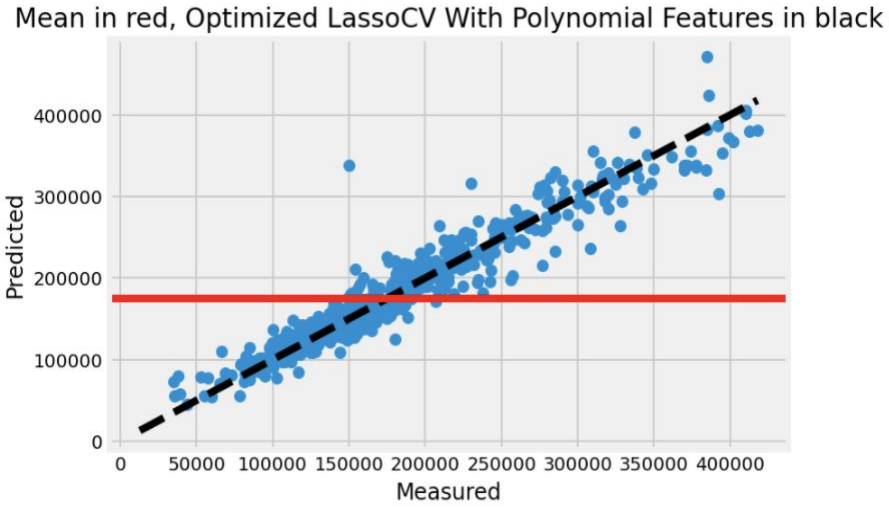
Coy B's Model



Prediction buffer
$22,673.49

Predictive power
past *clients* 89%
prospective *clients* 90%

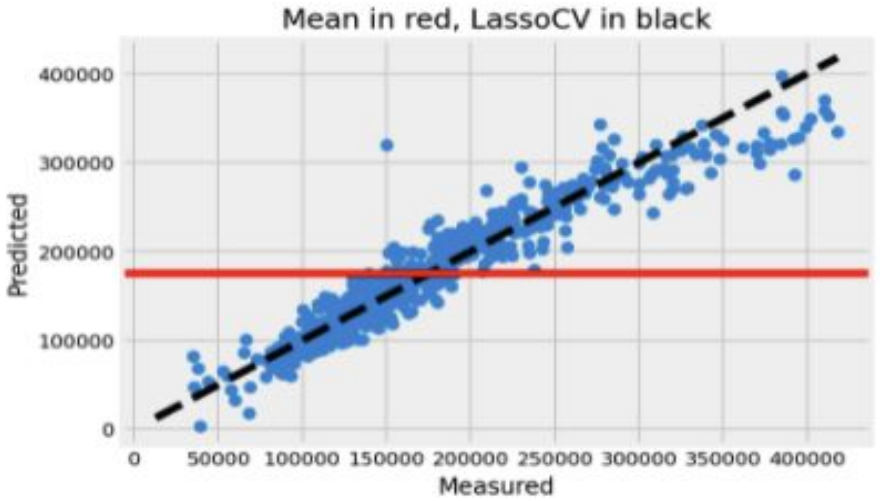
Proprata's Model



Prediction buffer
$20,253.00

Predictive power
past *clients* 92%
prospective *clients* 93%

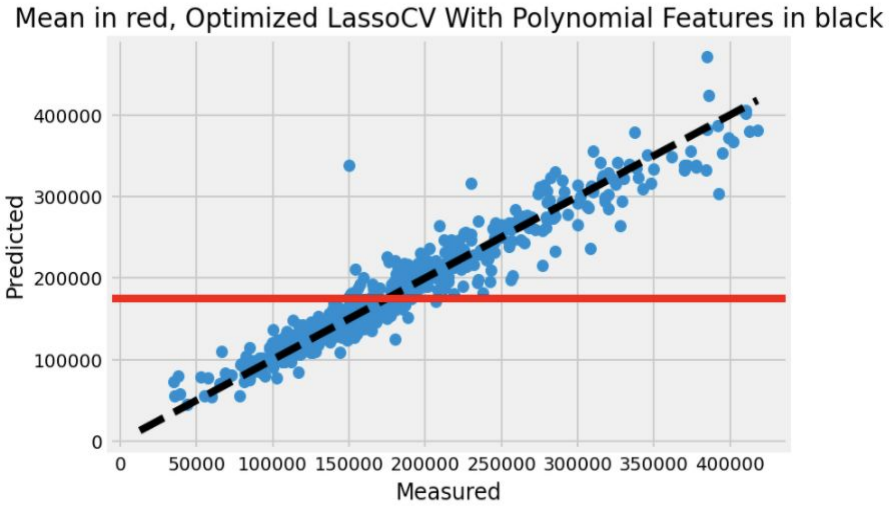
Coy C's Model



Prediction buffer
$22,685.97

Predictive power
past *clients* 89%
prospective *clients* 90%

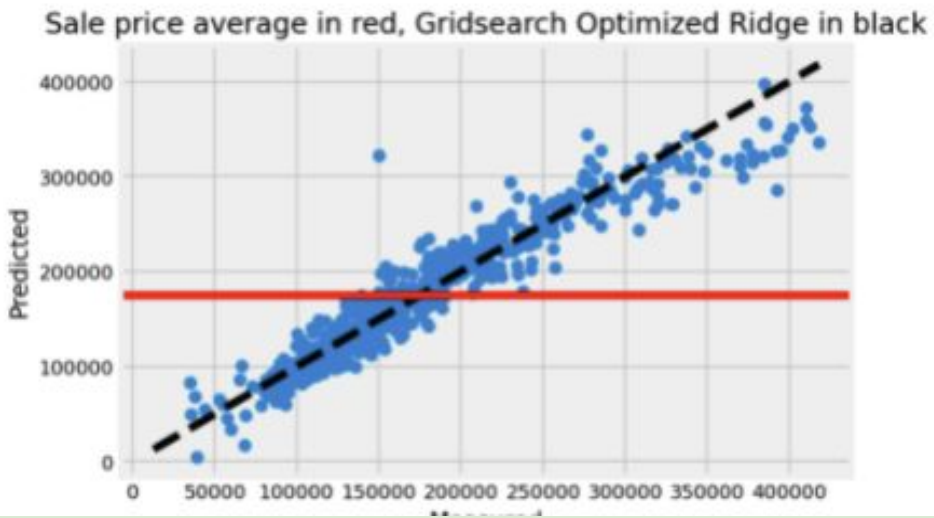
Proprata's Model



Prediction buffer
$20,253.00

Predictive power
past *clients* 92%
prospective *clients* 93%

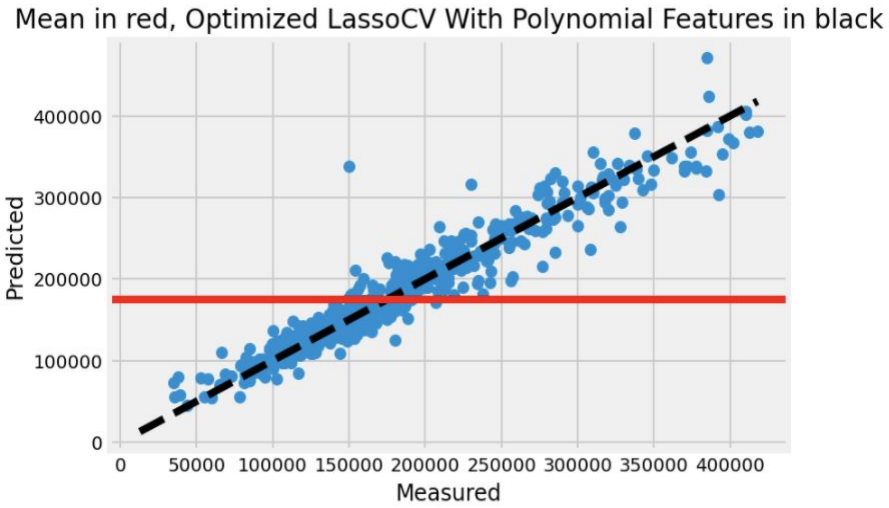
Coy D's Model



Prediction buffer
$22,676.54

Predictive power
past *clients* 89%
prospective *clients* 90%

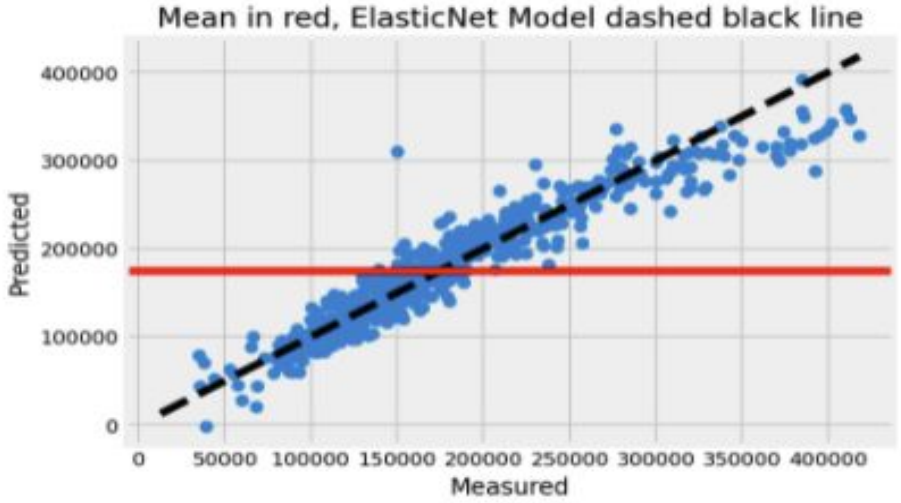
Proprata's Model



Prediction buffer
$20,253.00

Predictive power
past *clients* 92%
prospective *clients* 93%

Coy E's Model



Prediction buffer
$22,752.07

Predictive power
past *clients* 89%
prospective *clients* 90%

Proprata's
Optimized LassoCV with Polynomial Features Model

$20,253.00

Prediction buffer

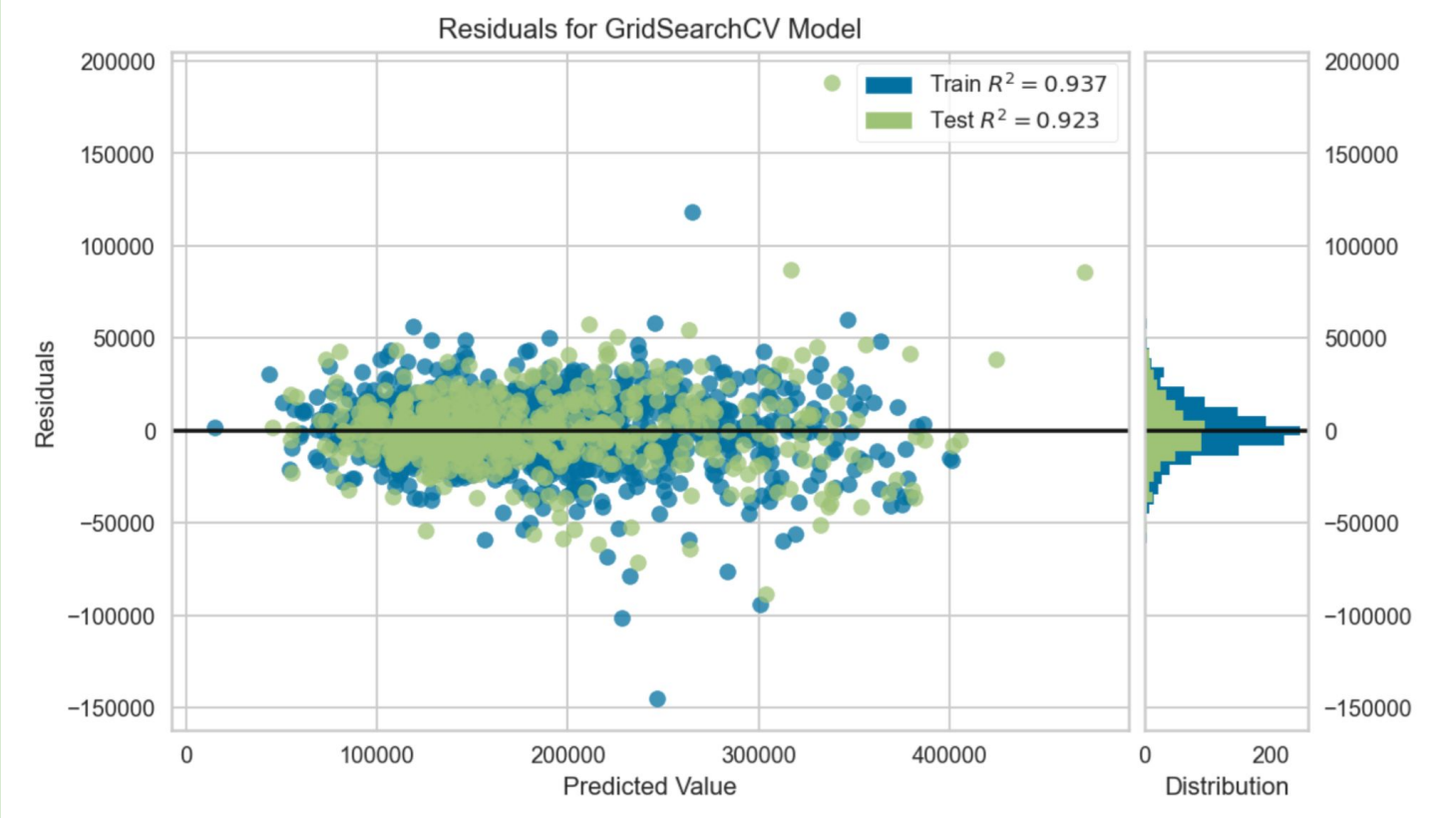
92%

Predictive power : ***past*** clients

93%

Predictive power : ***prospective*** clients

The Predictive Power of Proprata's Model



19591.74

Kaggle Submission Score

04

Conclusion and Recommendations

What we have to offer

Top 20 Features

Rank	Features	Rank	Features
1	Rooms above size	11	External features
2	Single storey	12	Garage area
3	Basement ceiling	13	Ground living area
4	Outside space	14	Lot size
5	Neighbourhood quality	15	Kitchen quality
6	Local features	16	Basement size
7	Fireplace quality	17	Duplex building
8	Paved driveway	18	Remodeled
9	Lot frontage	19	Heating quality
10	Floor size	20	Basement quality

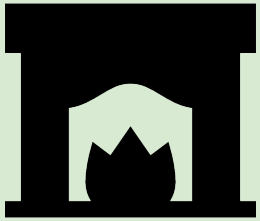
Features to buy into:

Rooms above Size

Single Storey

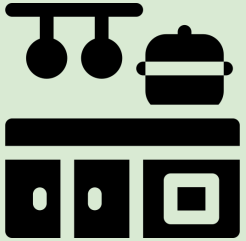
Neighborhood Quality

Features to reno:



Fireplace Quality

Paved Driveway



Kitchen Quality

Insignificant features:

Roof Quality

Age

Bath Rooms

Customised recommendations

Case Study

Seller X

Budget constraints

Consultation

Went through our winning formula and took our advice.

Renovation

Upgraded his fireplace for

$280

Result

Predicted increase in saleprice

$2838

TODAY only.

“Real Estate cannot be lost or stolen.

Nor can it be carried away.

Purchased with common sense,

paid for in full,

and managed with reasonable care,

it is about the safest investment in the world.”

—Franklin D’ Roosevelt

Thank you

List of references

- https://www.bostonfed.org/-/media/Documents/conference/36/conf36g.pdf?la=en
- https://ideas.repec.org/a/tpr/restat/v90y2008i4p599-611.html
- https://fred.stlouisfed.org/series/ASPUS
- https://www.statista.com/statistics/226144/us-existing-home-sales/
- https://www.propertyguru.com.sg/property-guides/property-flipping-singapore-boomer-millennial-28396
- https://en.wikipedia.org/wiki/United_States_housing_bubble
- https://www.investopedia.com/articles/mortgages-real-estate/08/house-flip.asp
- https://www.investopedia.com/terms/f/flipping.asp
- https://www.analyticsvidhya.com/blog/2021/05/yellowbrick-visualization-for-model-predictions/
- https://medium.com/mlearning-ai/a-thorough-dive-into-the-ames-iowa-housing-dataset-part-1-of-5-7205093a5a53
- http://jse.amstat.org/v19n3/decock/DataDocumentation.txt
- https://www.youtube.com/watch?v=VSeGseoJsNA
- https://www.homeyou.com/ia/chimney-repair-ames-costs
- https://newsilver.com/flipping-houses-in/ames-ia/
- https://newsilver.com/the-lender/flipping-houses-in-iowa/
- https://newsilver.com/the-lender/how-to-calculate-arv-simple-guide/
- https://web.missouri.edu/segerti/capstone/northcraft_neale.pdf
- https://www.trulia.com/blog/2-most-important-rooms-in-a-house/