

Real Estate Investment Proposal

...

August 21, 2020

Summary

Goal

- Define the Problem
- Data Understanding

Predictive Analysis

- Model Performance
- Prediction

Conclusion

- Recommendations
 - Areas for Growth
-

Goal:

Predict top 5 zip codes in United States for real estate firm to invest in.

Define the Problem

Rank “top 5” zip codes by Return on Investment (ROI) after filtering for historical fluctuations:

$$\text{ROI} = \frac{\text{Predicted Price Increase}}{\text{Initial Investment}}$$

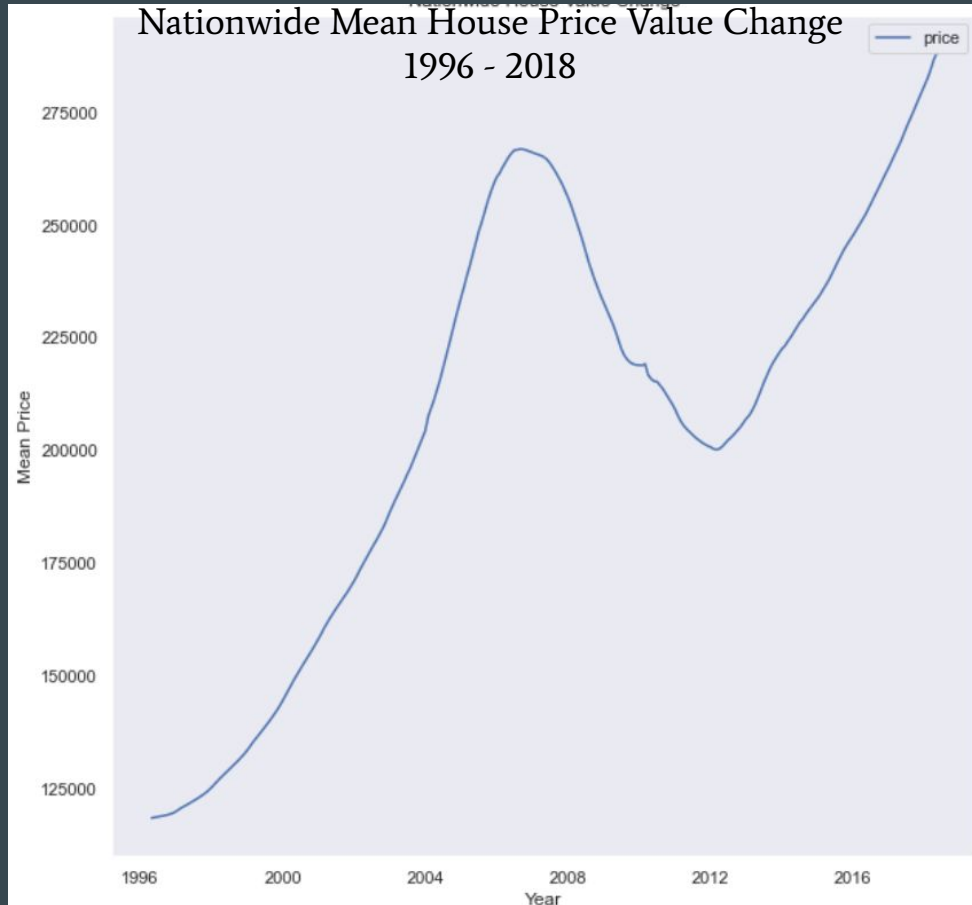
Data Understanding

- Raw data from Zillow
 - 1996 - 2018, Zip Codes across US
 - Limited to post-2008 market crash
 - Improves the model performance, since (hopefully) it was an anomaly
 - Narrowed focus: 60 zip codes w/ largest ROI in 5-year rolling window
 - Selected smaller subset of this initial filtering to enable faster model training
-

Data Understanding

Time Series Analysis:

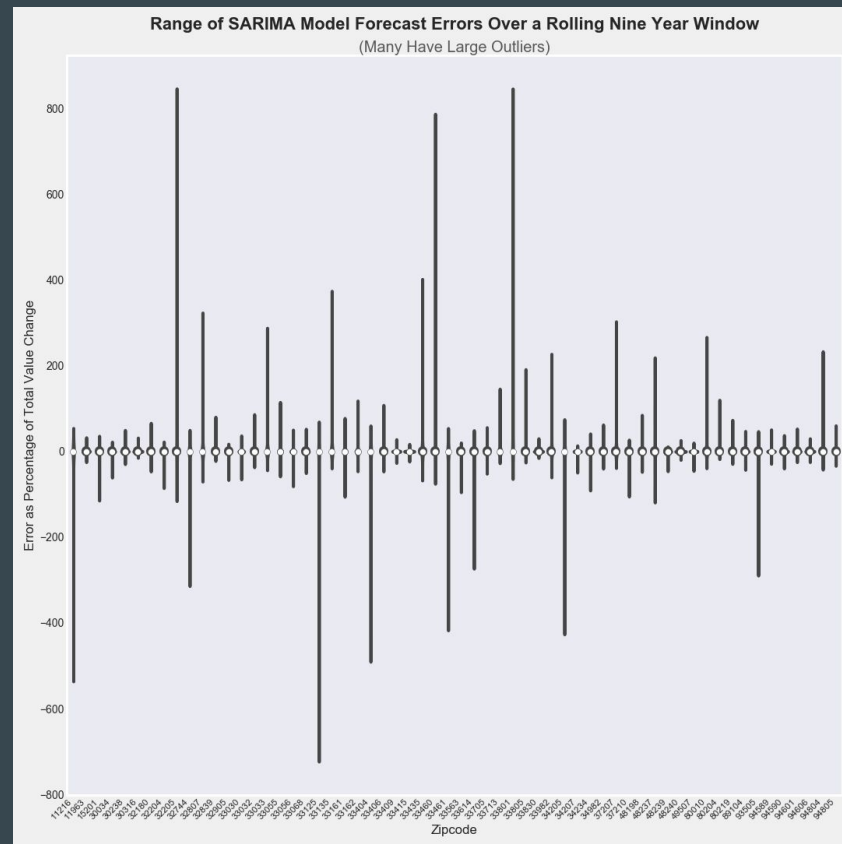
- Eliminates noise in data (trend, seasonality, autocorrelation, stationarity)
- Enables value prediction



Predictive Analysis

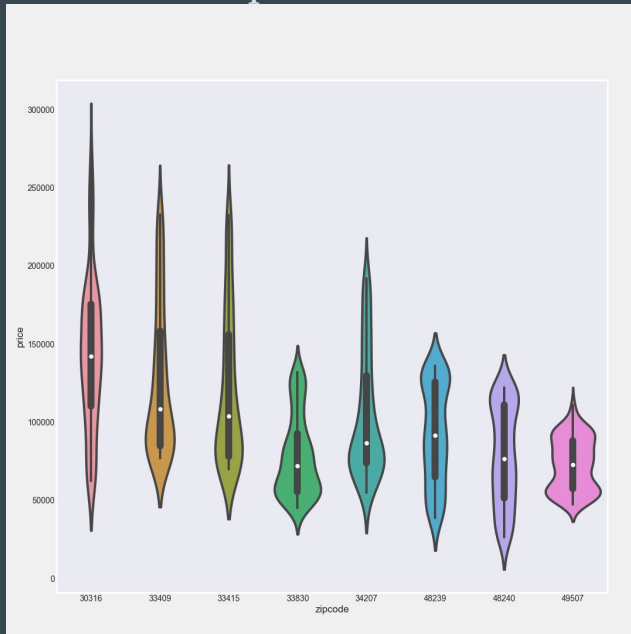
Model Performance

- Large errors in several zip codes
- Limited to 10 w/ smallest range of error over time

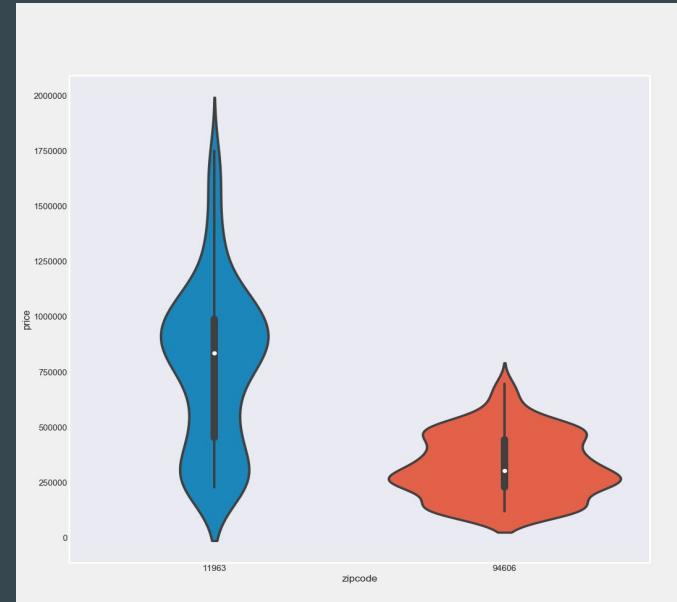


Model Performance: Two Investment Points

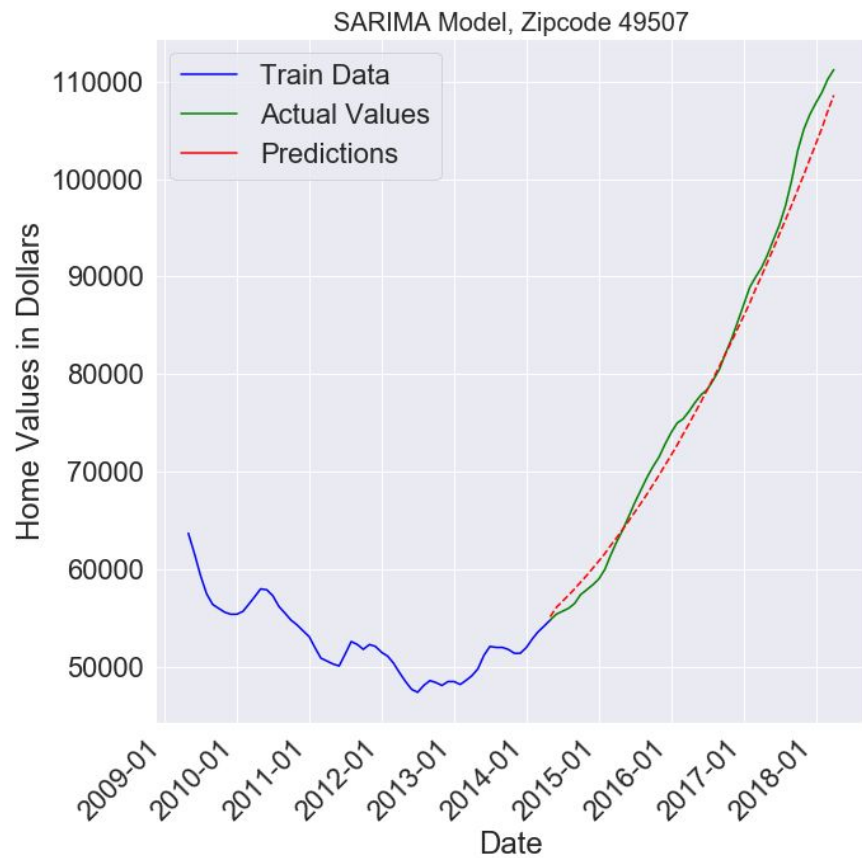
Lower Median Price
(lower capital investment)



Higher Median Price
(higher capital investment)



Model Performance



Conclusion

Recommendations: Zip Codes to Invest In

Low Medians

	ROI	RMSE as Percentage of Total Change
34207	108.78%	6.44%
33415	108.41%	7.66%
33830	100.4%	8.07%
49507	100.34%	3.37%
33409	87.44%	3.42%
48240	84.45%	6.11%
30316	79.23%	4.76%
48239	70.54%	5.13%

High Medians

	ROI	RMSE as Percentage of Total Change
94606	91.75%	3.02%
11963	129.71%	9.5%

Zip Codes beginning w/ 3 are mostly Florida

Zip Codes beginning w/ 4 are mostly Michigan

Outlier, 11963, is in New York State

Areas for Growth

Include Zillow Rent Data

- Predicted rent would provide another way to predict return on investment

Extend time frame for analysis

- With more data, predict longer term frames of reference

Improve Model

- Train model on nationwide data for more precise “drill-down” ability

Thank You!



Jake Miller Brooks



Alexander Newton

