Determining Optimal Real Estate Investment Opportunity

●●●

Using Time Series Modeling to Analyze Zillow Data

By Orin Conn

Overview

1. Task Definition & Data
2. Methodology for Narrowing Down Selection of ZIP codes
3. Visualizing Top ROI Contenders
4. Time-Series Modeling Our #1 Pick
5. Forecasting Our Top 5 ZIP codes
6. Conclusion and Further Recommendations

Task:
Find the top 5 best ZIP codes to invest in.

Data Utilized:

- Zillow Research Dataset
  - +14,000 ZIP codes
  - Median House Price/Month
    - Data from 1996 - 2018

Methodology
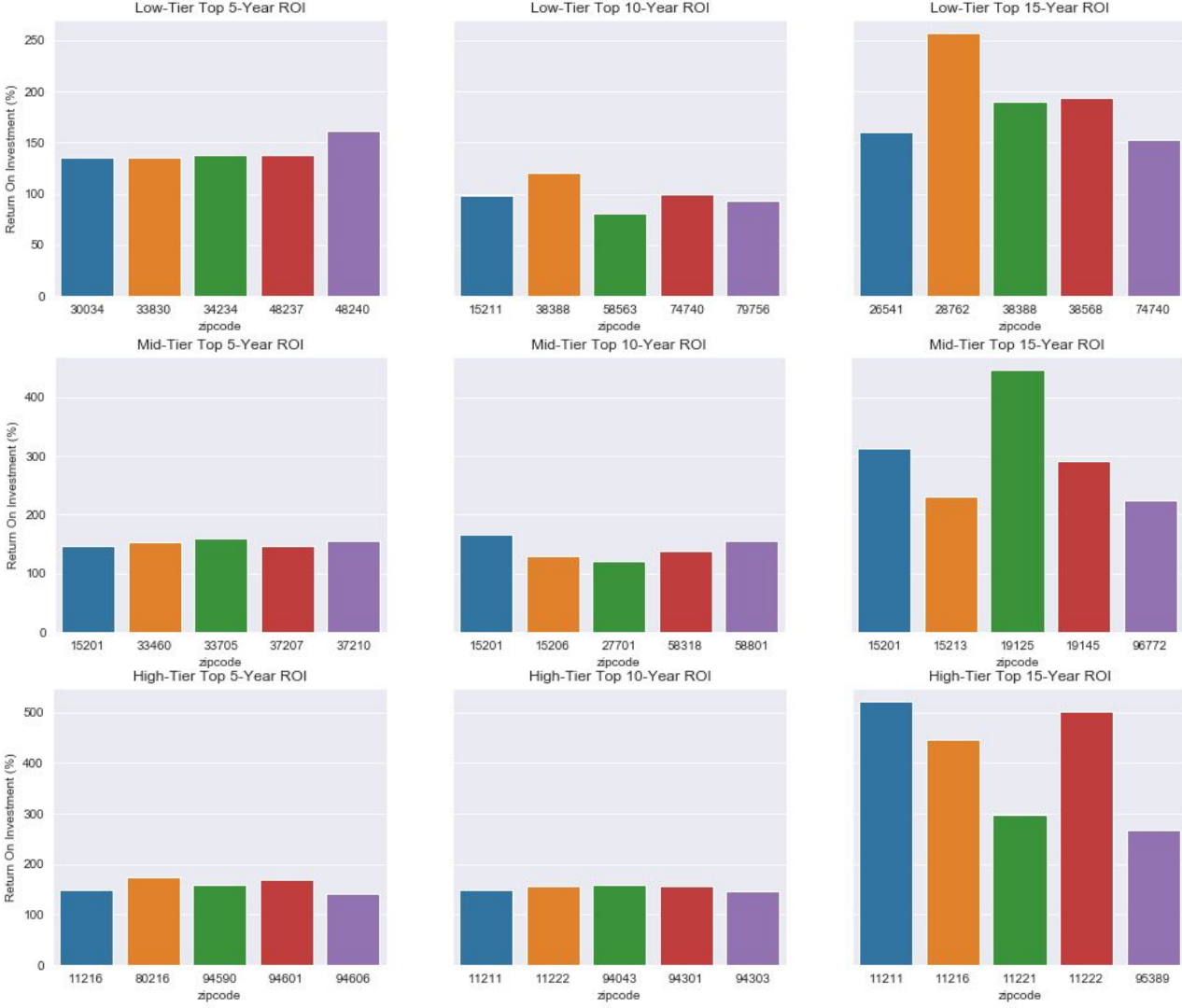
1. Define "Best" Investment

- Low Risk
  - Low volatility in price
- Upward trend
- Highest Forecasted ROI
  - 5-yr, 10-yr, 15-yr
  - Low RMSE

2. Filtering Data

- Split data into 3 equal price percentile tiers
  - < $151k
  - $151k - $263k
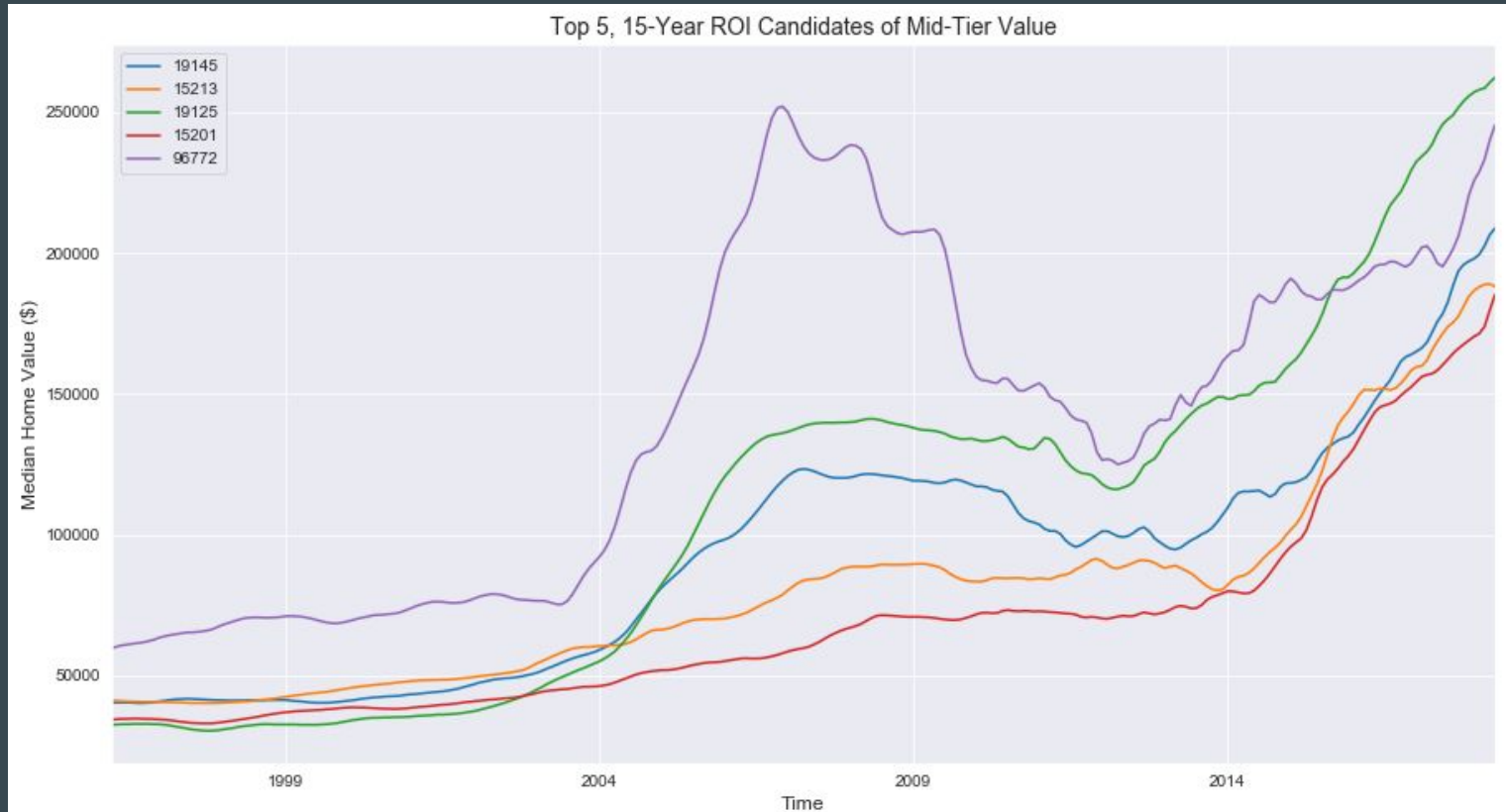 - > $263k
    - Best 5y, 10y, 15y ROI for each tier

3. Selecting Lowest Risk

- Visualize trends of top 5 ROIs for each tier and each subgroup
- Select least visually volatile for time series modeling & forecasting
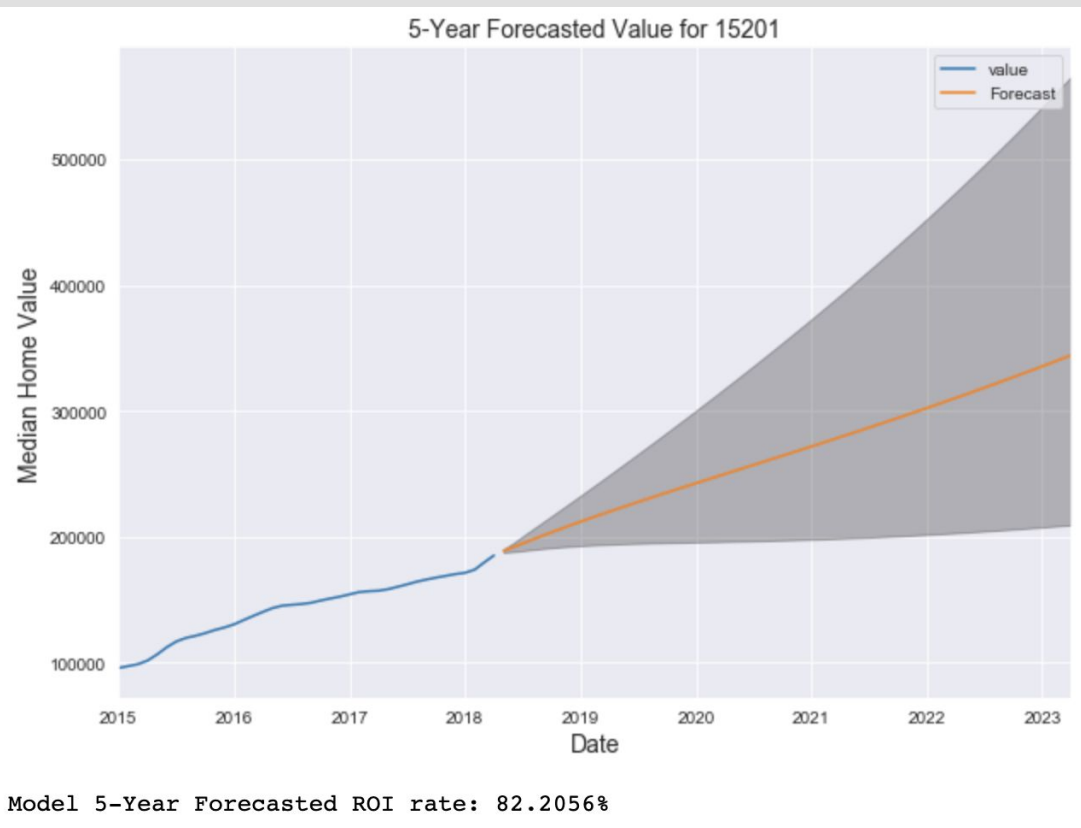
TOP ROI COMPARISON

SELECTING #1 - 15201

Top 5, 15-Year ROI Candidates of Mid-Tier Value

15201

ZIP 15201 Annual Price Distribution, 2007 - 2017

Best Time Series Model:

- Log transform of data
- No seasonality
- RMSE: 921.98

5-Year Forecasted Value for 15201

Model 5-Year Forecasted ROI rate: 82.2056%

15201 (Mid-Tier Investment)
Forecasted 5-Year ROI:

82.20%

Top 5 ZIP Codes

	CURRENT MEDIAN VALUE	RMSE	5-yr Forecasted ROI
74740:	$ 80.1k	297.3	+28.59%
15201:	$ 185.2k	921.9	+82.20%
15206:	$ 191.1k	1,151.5	+53.62%
11222:	$ 1.72M	11,193.8	+81.46%
94043:	$ 2.07M	7,539.3	+71.32%

Conclusions:

- Best Model:
    - No Seasonality
    - Unique parameters for each ZIP

- Best ZIP codes:
    - High ROI
    - Resistance to '07/'08 financial crisis
        - Lower Risk

Further Recommendations:

- More time to quantify volatility
- More computing power for greater range of model parameters
- Consider effects of COVID on housing market
    - More recent dataset

Thank You!