

# OpenDoor

Market Recommendations for 2021

**Alexis Kedo**

---

<https://github.com/alexiskedo>

[alexiskedo@gmail.com](mailto:alexiskedo@gmail.com)



# Business Problem



## Room to Grow

Opendoor is currently in 27 markets and wants to capitalize on the favorable real estate market.



## Unfamiliar Territory

Opendoor needs to adapt to the shifting demographics brought about by the pandemic.



## New Opportunities

I sought to identify the most promising areas for investment in high-growth areas.

# Data

A combination of population data from the U.S. Census and Zillow real estate data

14,723

Unique Zipcodes

---

June 2009 - April 2018

Timeframe used to predict 2021 Return on Investment (ROI)

---

2010

Year of most recently available U.S. Census data

# Methods

Before predicting, zipcodes were split into three groups and sorted based on population.

The top 10 zipcodes with the best average ROI were then selected.

## ◆ **Very Dense/Urban**

- Zipcodes with > 100,000 residents
- 22-Year Average ROI between 98- 99%.

## ◆ **Dense/Mid-Size**

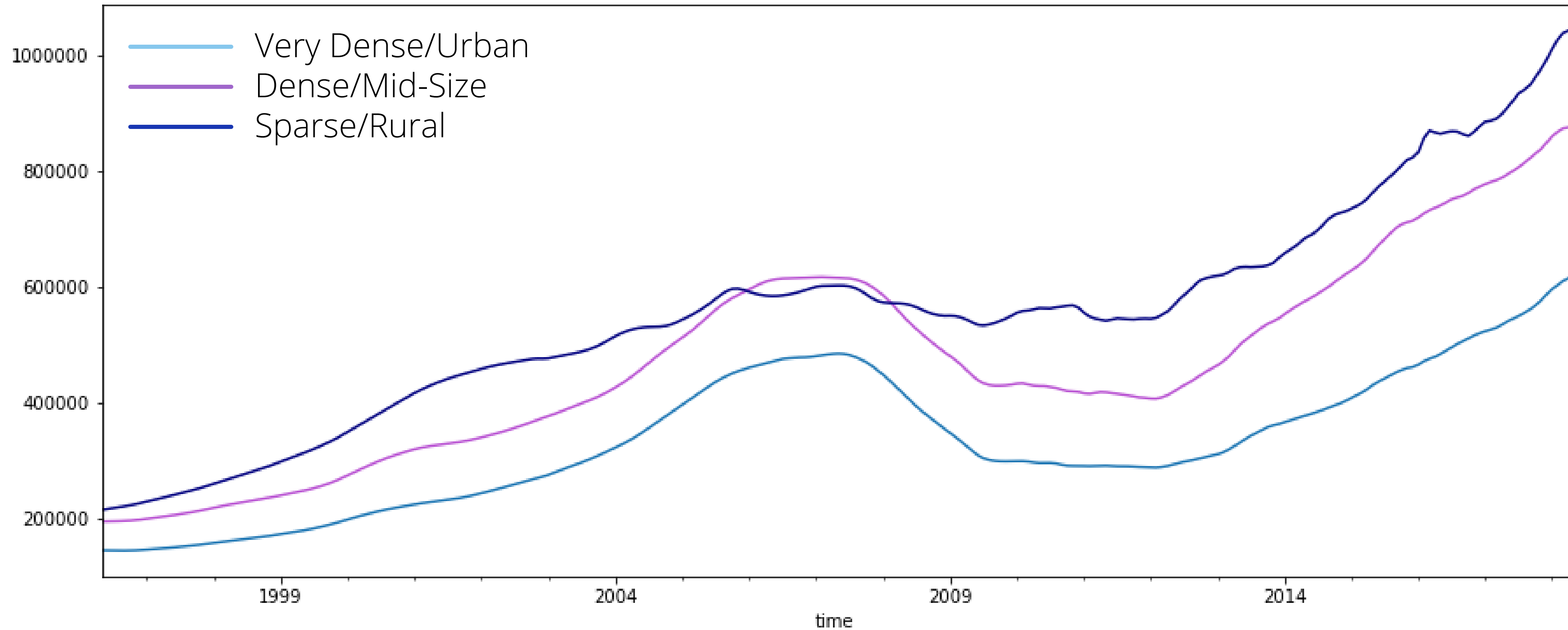
- Zipcodes with 25,000-99,000 residents
- 22-Year Average ROI between 98-99%

## ◆ **Not Dense/"Small"**

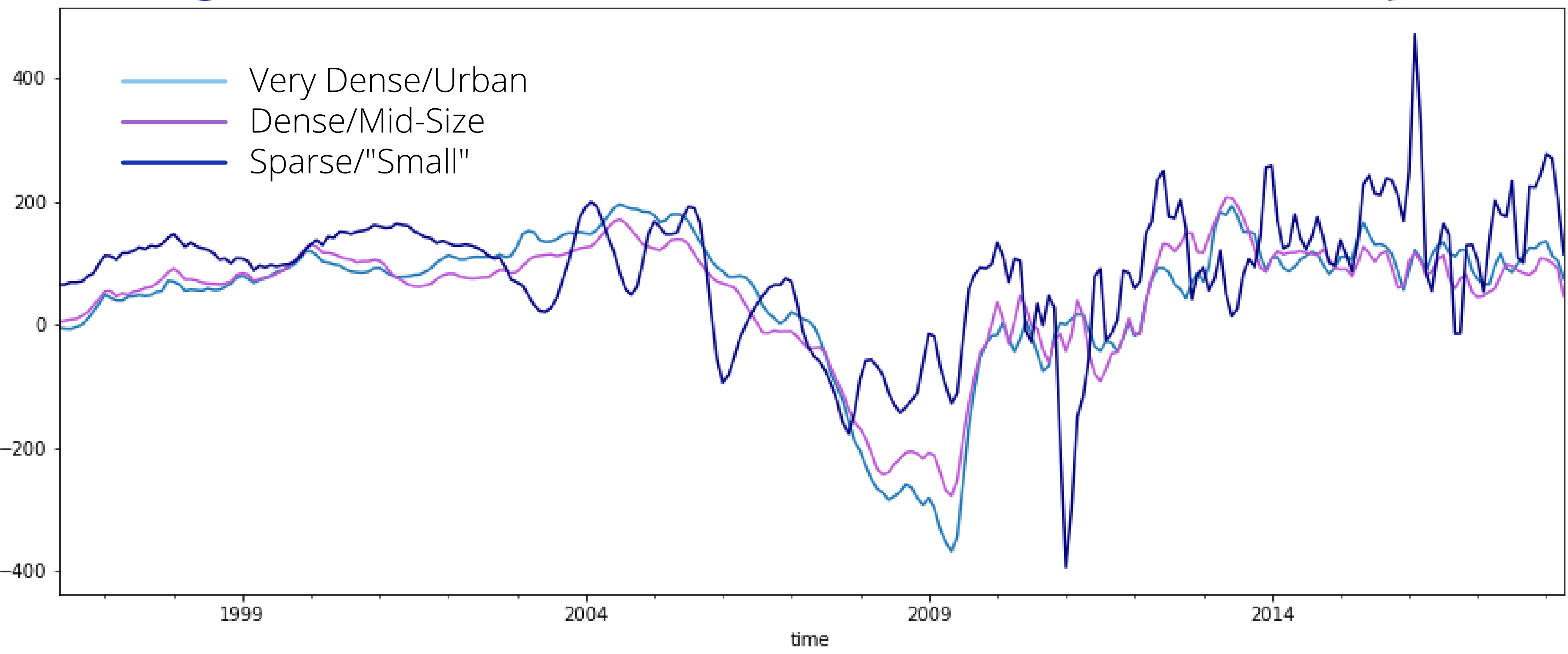
- Zipcodes with <25,000 residents
- 22-Year Average ROI between 91-99%



# Average Value over Time for Most Profitable Zipcodes



# Average Return on Investment for Most Profitable Zipcodes



# Workflow

1

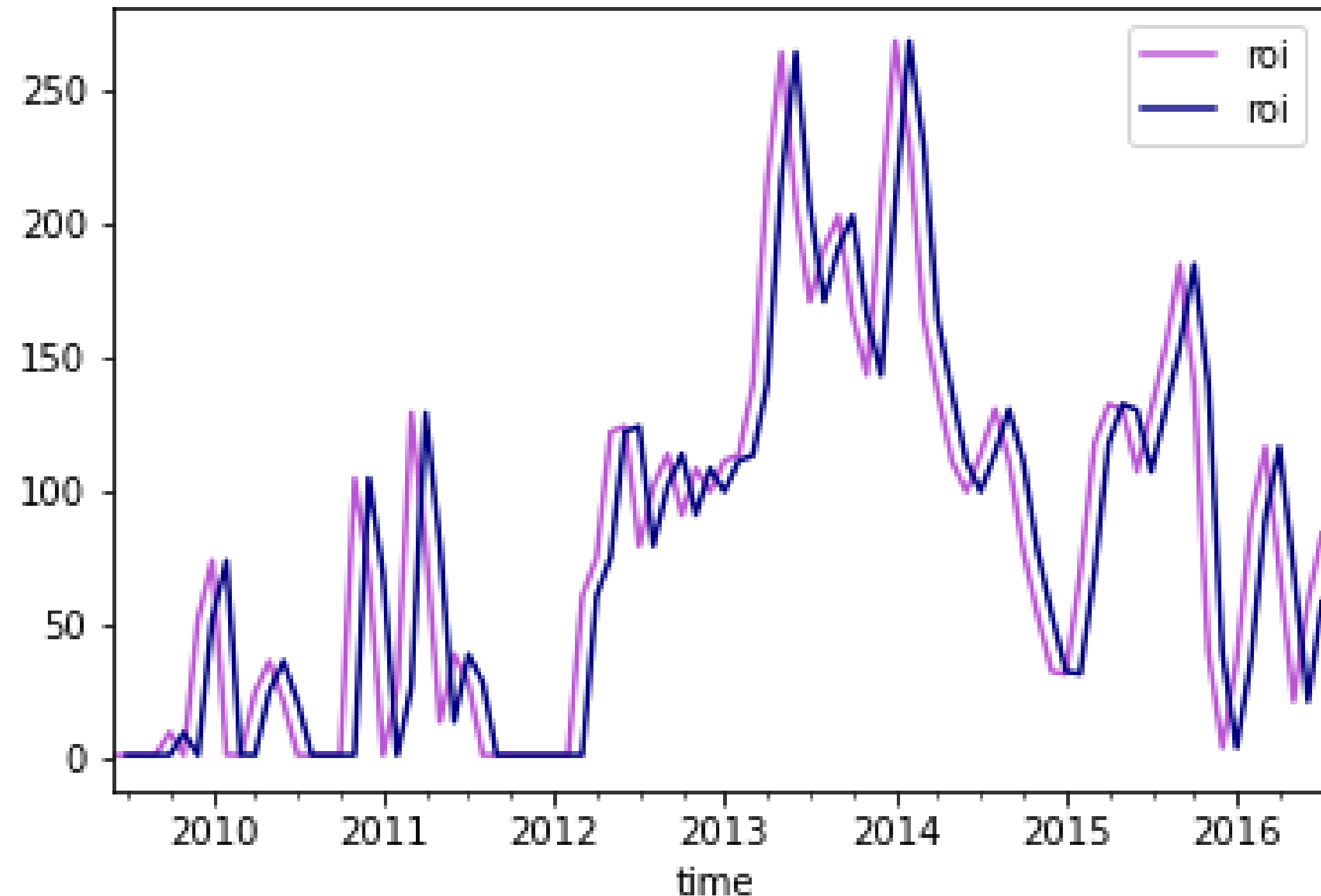
What happens if we simply predict the ROI for a representative zipcode based on the average ROI of the month prior?

The model is off by an average of 37.6% ROI.

Known Values  
Model Predictions



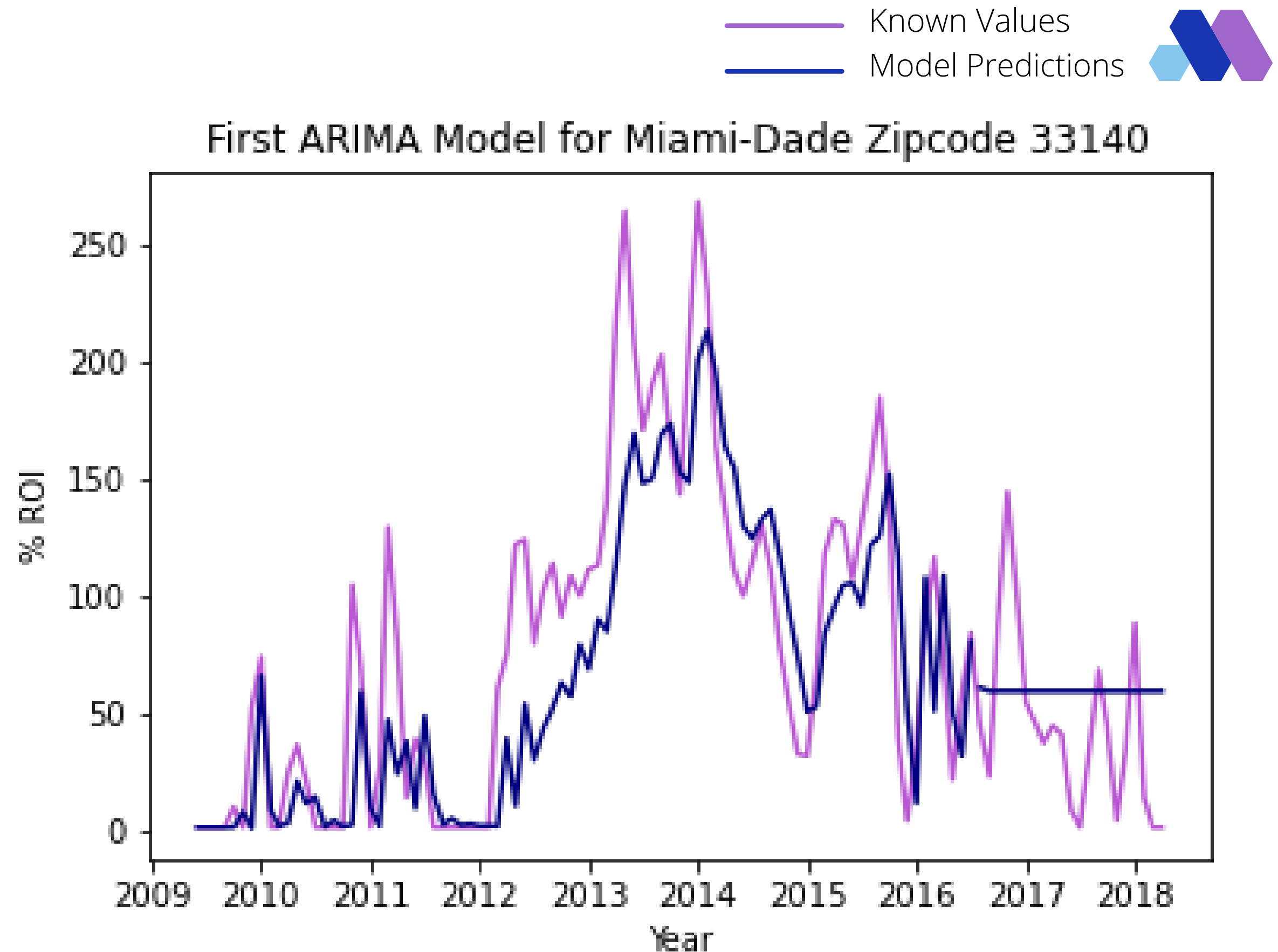
Random Walk for Miami-Dade Zipcode 33140



# Workflow

2

The next model attempted was an ARIMA model. This performed well on the train data but performed worse on the test data than the first model. It was off by almost 40% ROI.





# Workflow

3

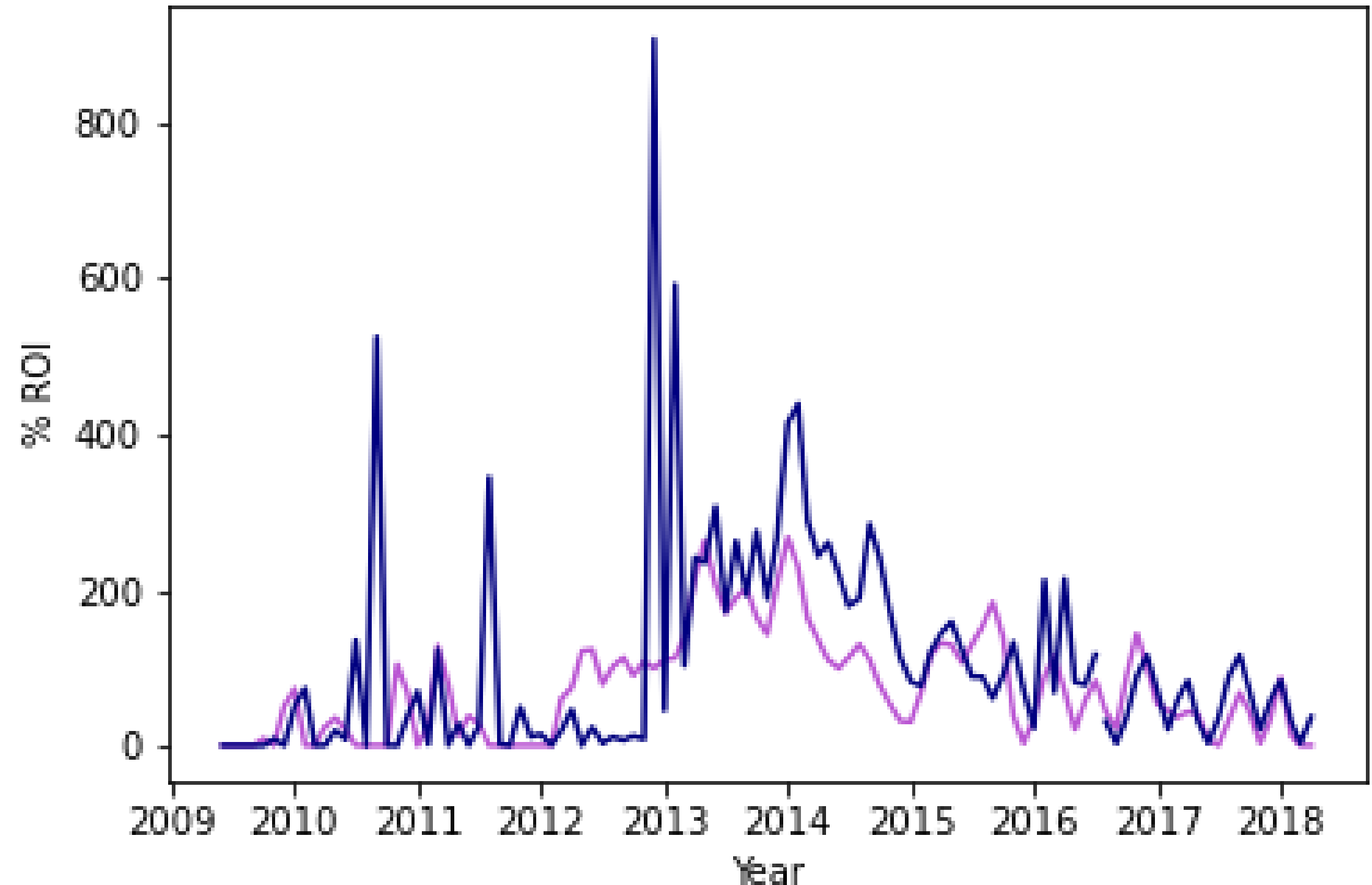
Refined model, including switching to a SARIMA model to take into account seasonality.

This performed much better on the testing data, and was off by about 30% ROI -- an improvement from the first model.

Known Values  
Model Predictions



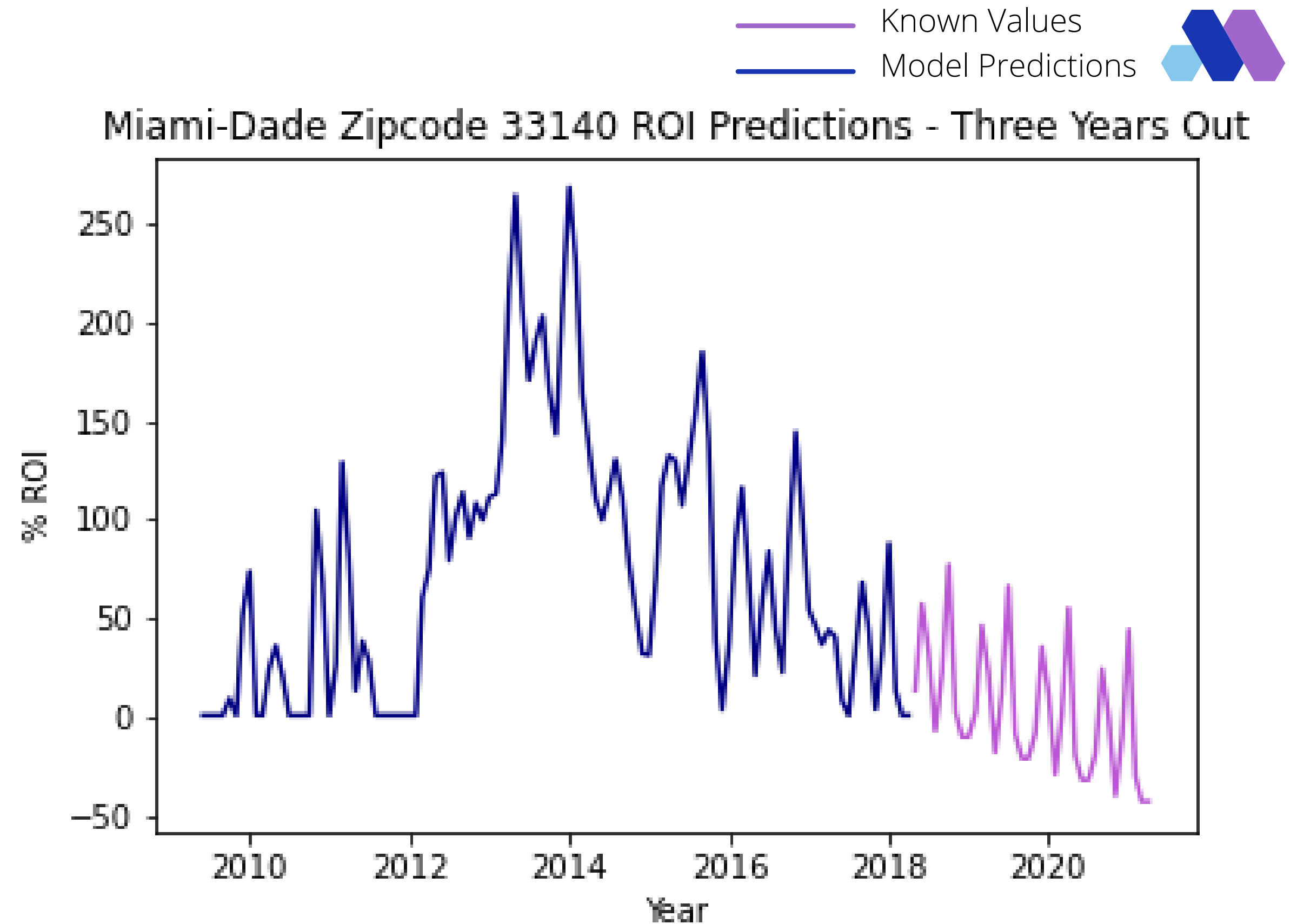
Final SARIMA Model for Miami-Dade Zipcode 33140



# Workflow

4

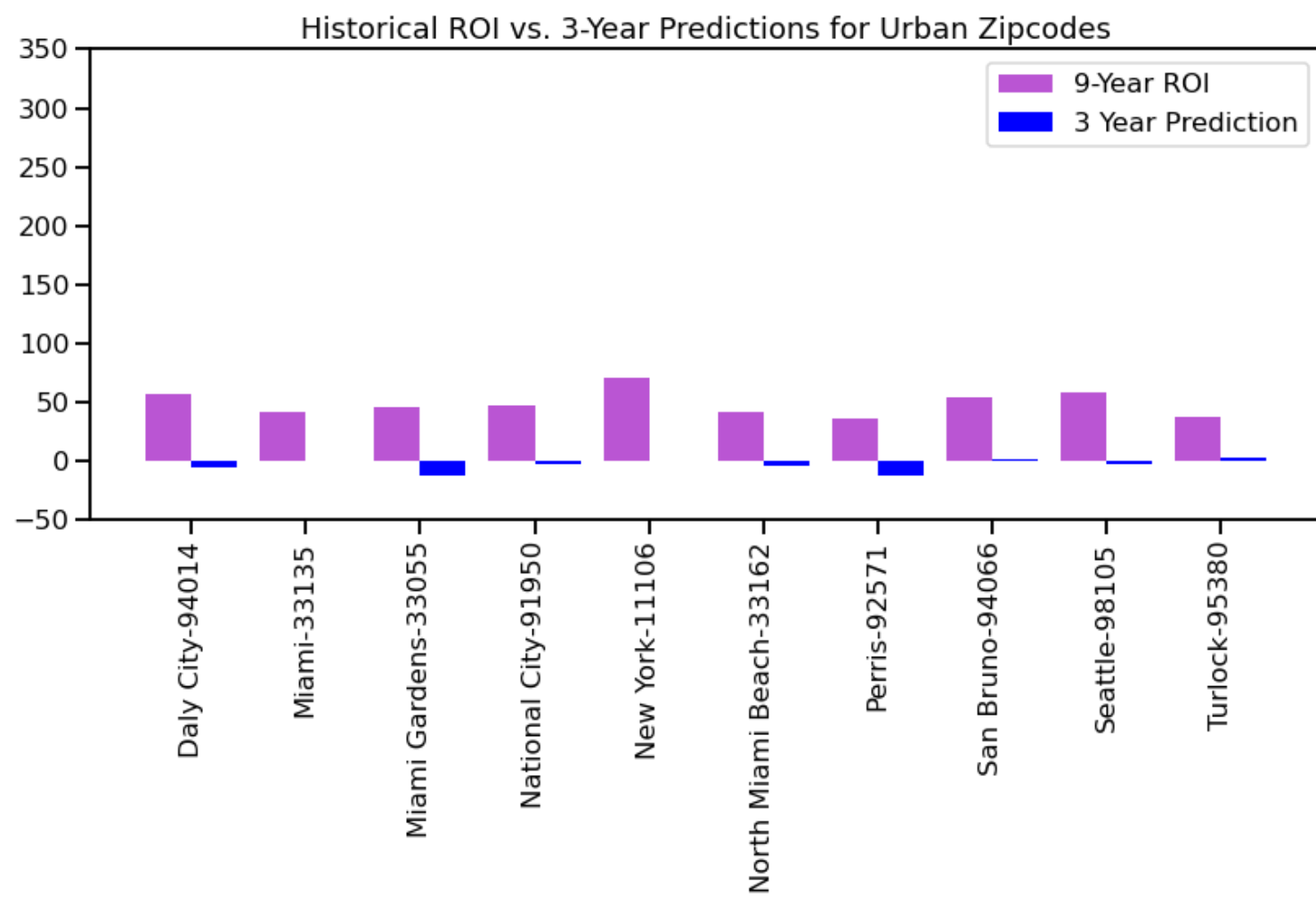
I then used the optimized model to predict the ROI three years into the future, first for the representative zipcode and then for all remaining zipcodes.



# Results

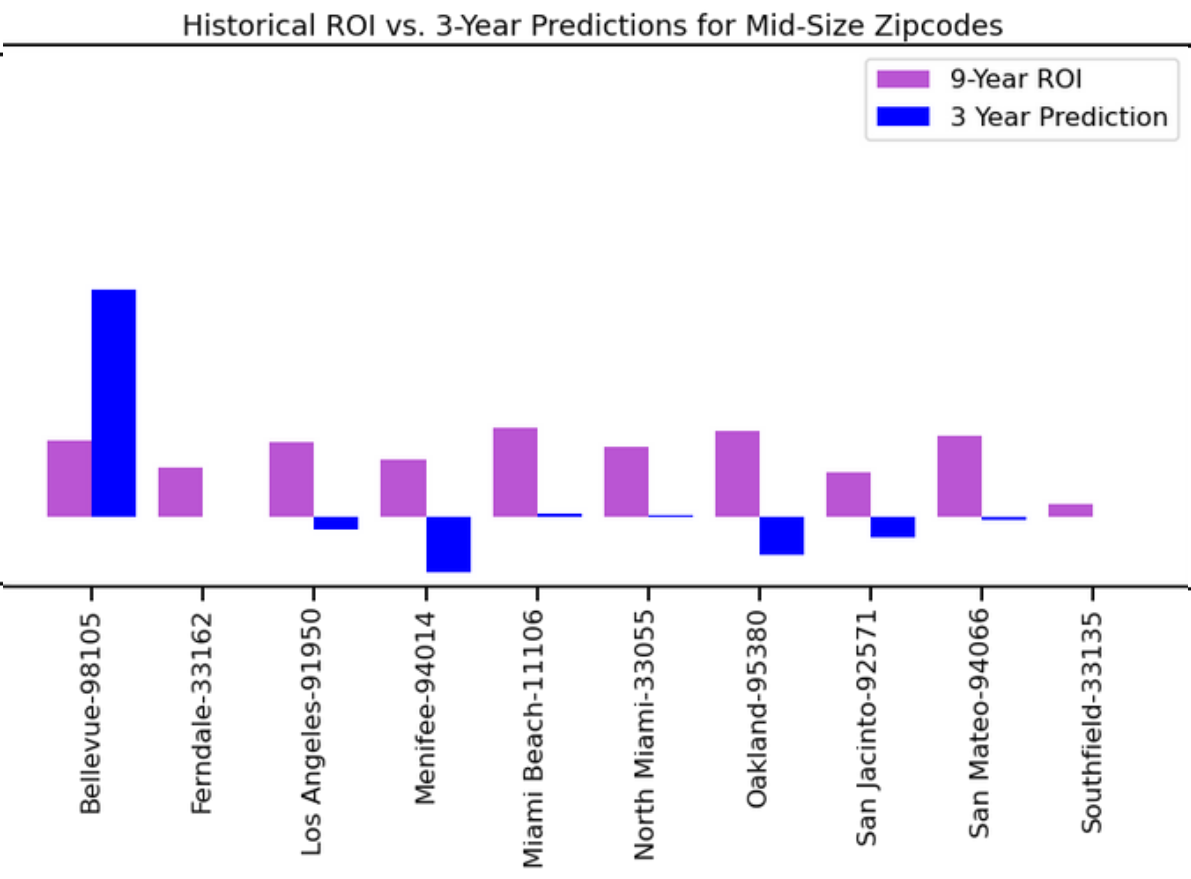
## Urban

-3.06% average predicted ROI



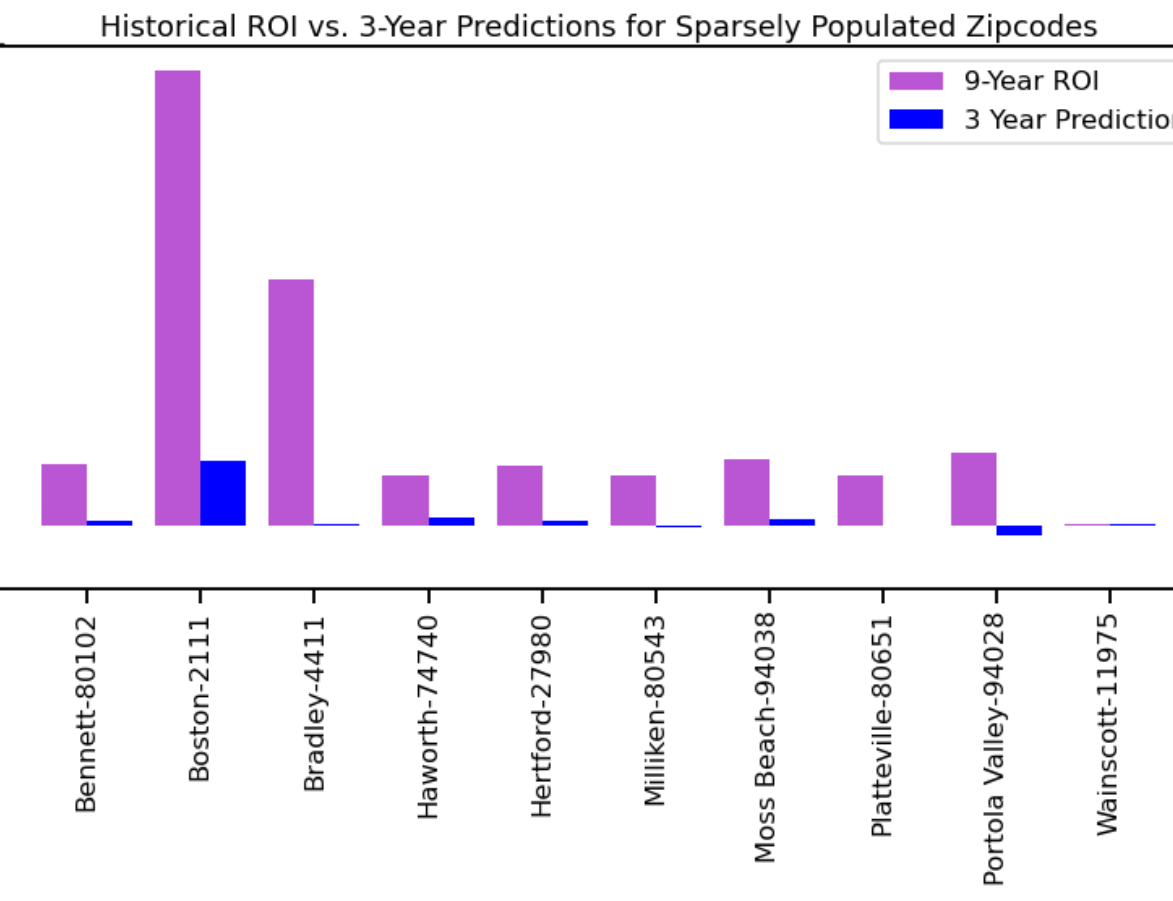
## Mid-Size

8.06% average predicted ROI



## Sparse

5.76% average predicted ROI



# Recommendations

## 1. Bellevue, WA (98105)

169% Predicted 3-Year ROI

## 2. Boston, MA (02111)

49% Predicted 3-Year ROI

## 3. Haworth, OK (74740)

6.24% Predicted 3-Year ROI

## 4. Moss Beach, CA (94038)

4.27% Predicted 3-Year ROI

## 5. Hertford, NC (27980)

3.74% Predicted 3-Year ROI

Urban areas are a riskier investment -- Opendoor should concentrate on high-potential mid-size cities .

Only one zipcode, 98008 -  
- located in outside of Bellevue, Washington--is predicted to be profitable within the next three years, with a predicted ROI of 169%.

# Next Steps

## Updated Data

Rerun models using population estimates from the 2020 Census.

## Differentiated Models

Since ROI for the "sparsely populated" category seems to follow a different pattern than the other two, we could optimize a model specifically for this group.



# Thank you!

---

[LinkedIn](#)

[GitHub](#)

[alexiskedo@gmail.com](mailto:alexiskedo@gmail.com)