# PHASE 4 PRESENTATION

Angelo Turri

Data

- 453 zip codes located in Michigan

- 76 months of data per zipcode (Jan 2012 – Apr 2018)

- 34,438 total rows of data

  - Each row represents the average home price in a zip code on a particular month
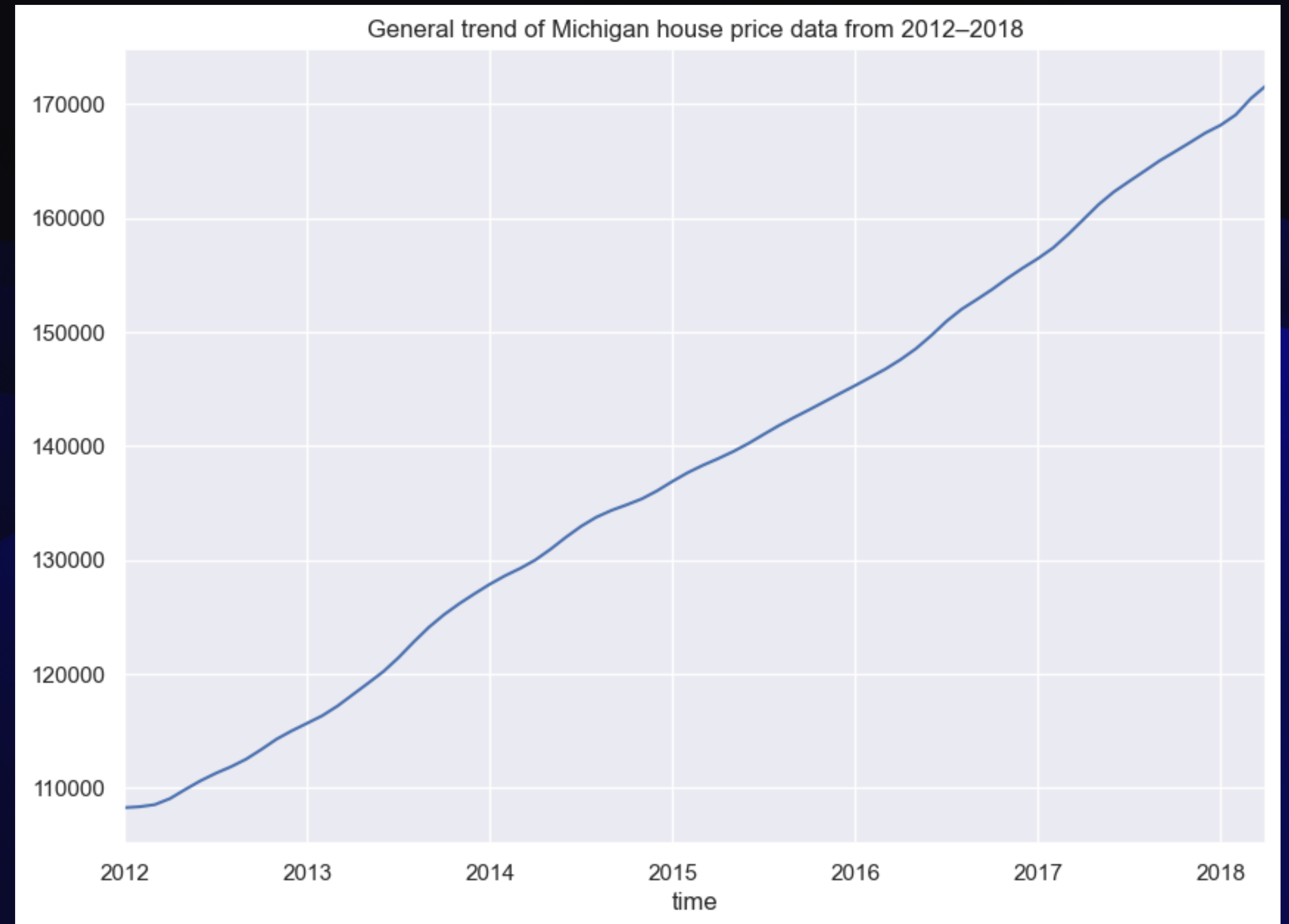
Distribution of target variable

- Average price around $130-140k

- Black line represents the average home price

Distribution of Michigan house prices between 2012 and 2018
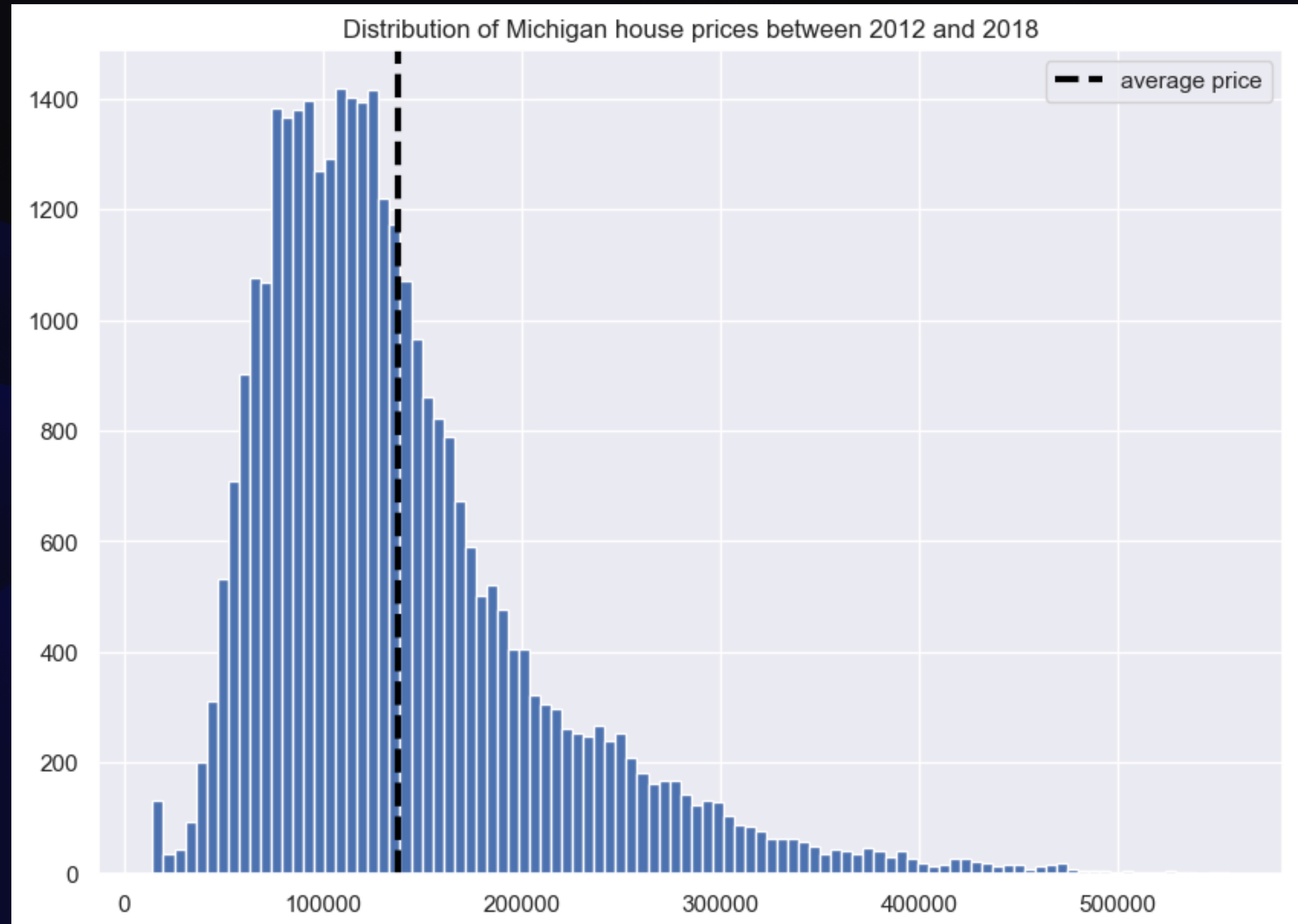
Trend of target variable

- Trend of house prices from 2012-2018

- Mostly linear

General trend of Michigan house price data from 2012–2018

Stakeholder

- A home-flipper in Michigan

- They are looking to buy properties, renovate them, and sell them after 3 months

- They want the five best zip-codes to invest in

  - Low volatility in price movement

  - Highest ROI after three months

Data preprocessing

- Only used data from after 2012; data before 2012 included the 2008 housing market crash.

- If our models learn from a single anomalous and highly unlikely occurrence, it could adversely impact their predictions.

Metrics

- Volatility score

  - The higher the volatility score, the more the price jumps around

  - We wanted a volatility score less than 5 because it seemed like an acceptable threshold of risk

- Root mean squared error

  - The average number of dollars that the model's predictions deviate from the true prices

- ROI (return on investment)

  - The three-month forecast for price divided by the latest recorded price.

  - Tells you how much your initial investment increased by.

Models used

- Iterative modeling process: start with a baseline, and make improvements from there.

- Final model (ARIMA model) – takes three parameters

  - p: each value regressed onto one or more of the previous values

  - d: how you difference your data (each value minus the value n values before it)

  - q: each value regressed onto one or more of the previous value's errors
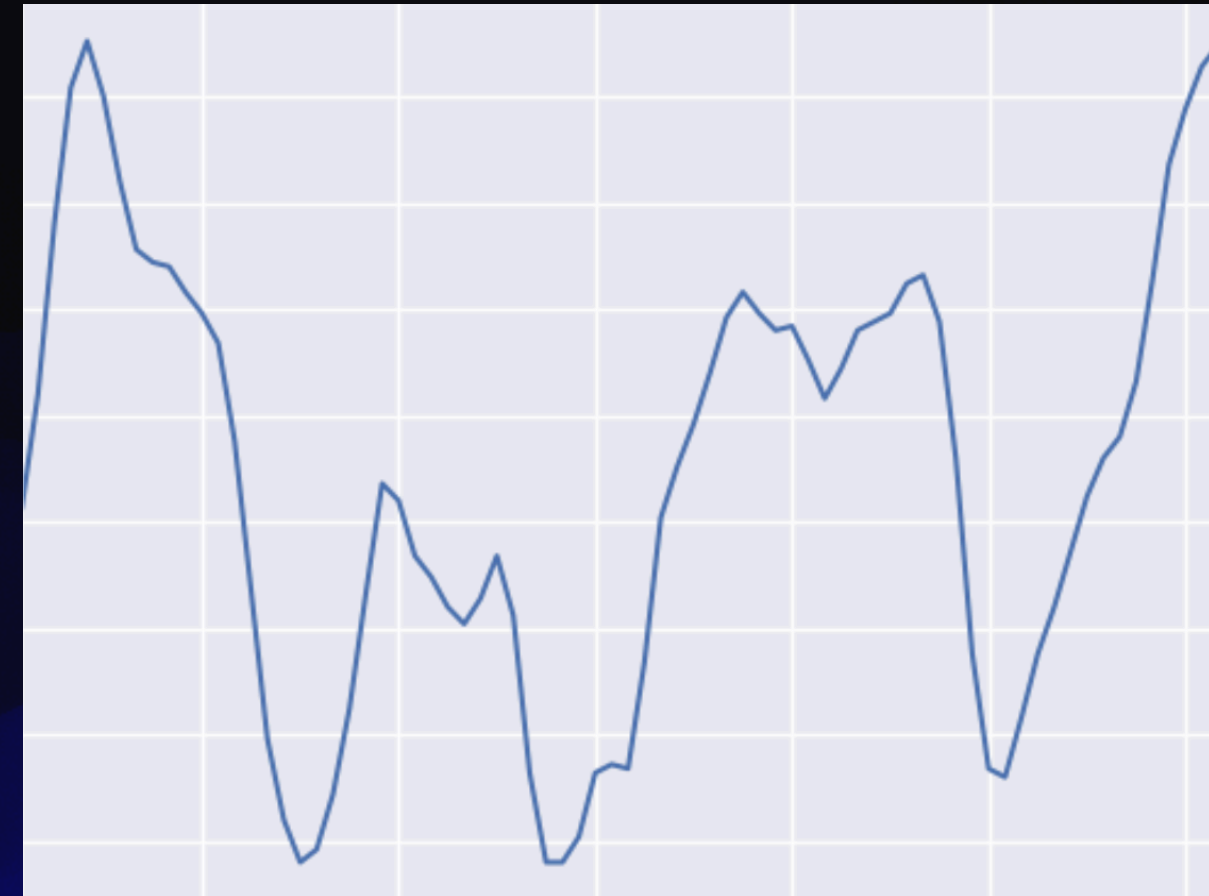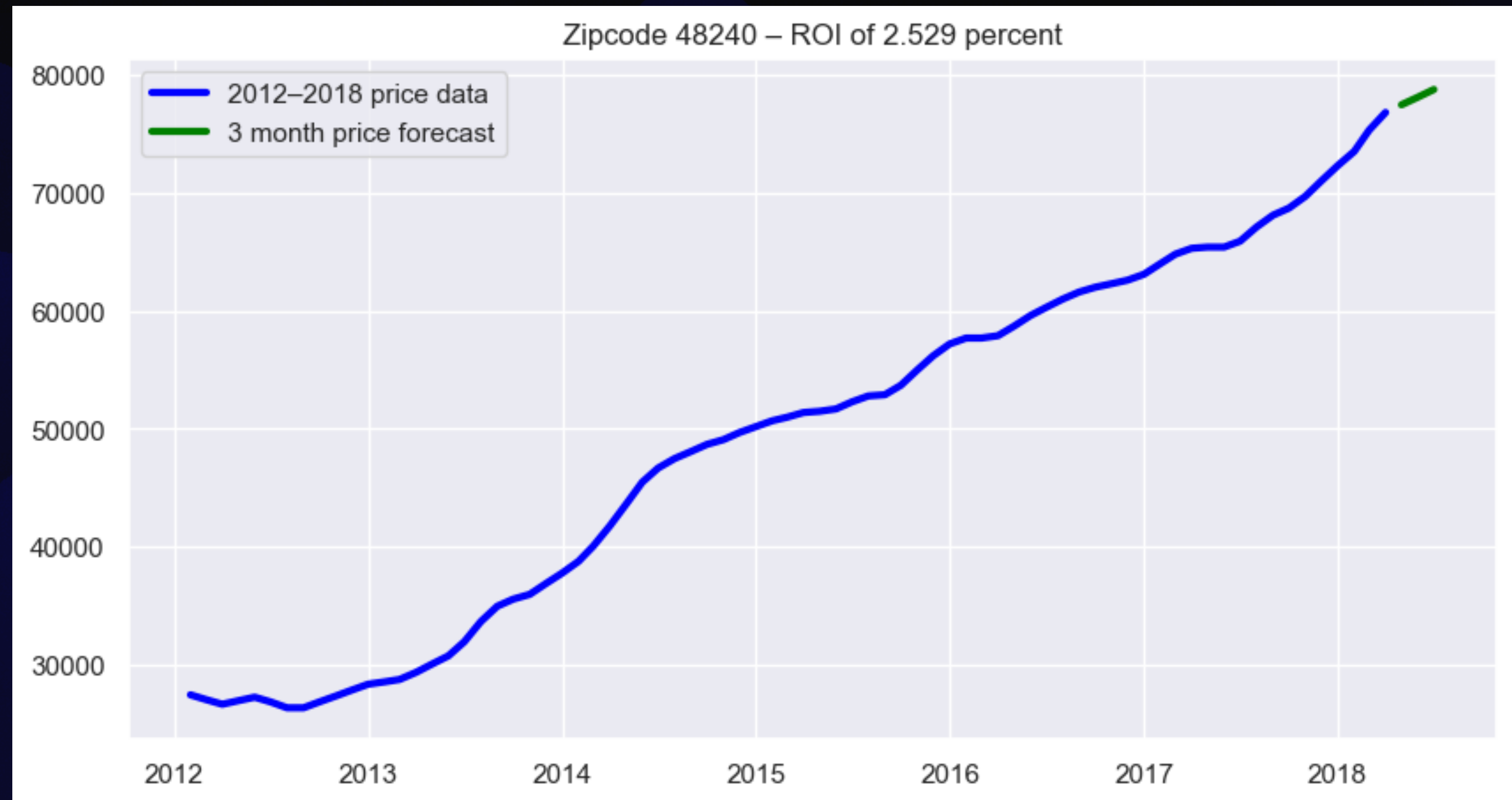
TRAIN RMSE

$950

TEST RMSE

$5,489

Leaving out volatility scores > 5

- These are three examples of the zip codes that we passed on.

Recommendation #1 – Zip code 48240

- Location: Redford, Wayne County MI

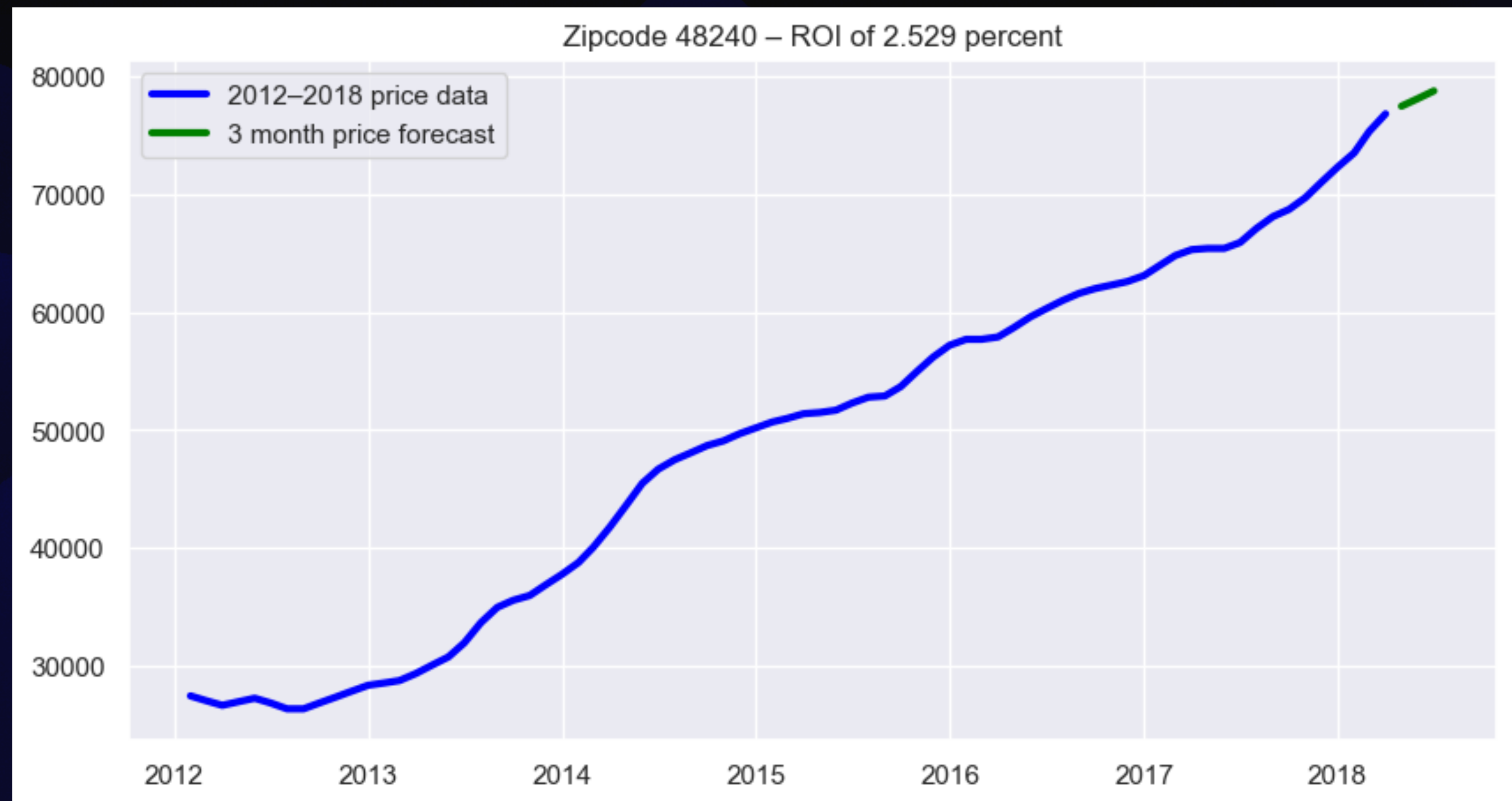- Expected ROI: 2.529%

- Final price prediction: $78,741

Zipcode 48240 – ROI of 2.529 percent

Legend: 2012–2018 price data (blue), 3 month price forecast (green)

Recommendation #2 – Zip code 48184

- Location: Wayne, Wayne County MI

- Expected ROI: 2.472%

- Final price prediction: $91,200

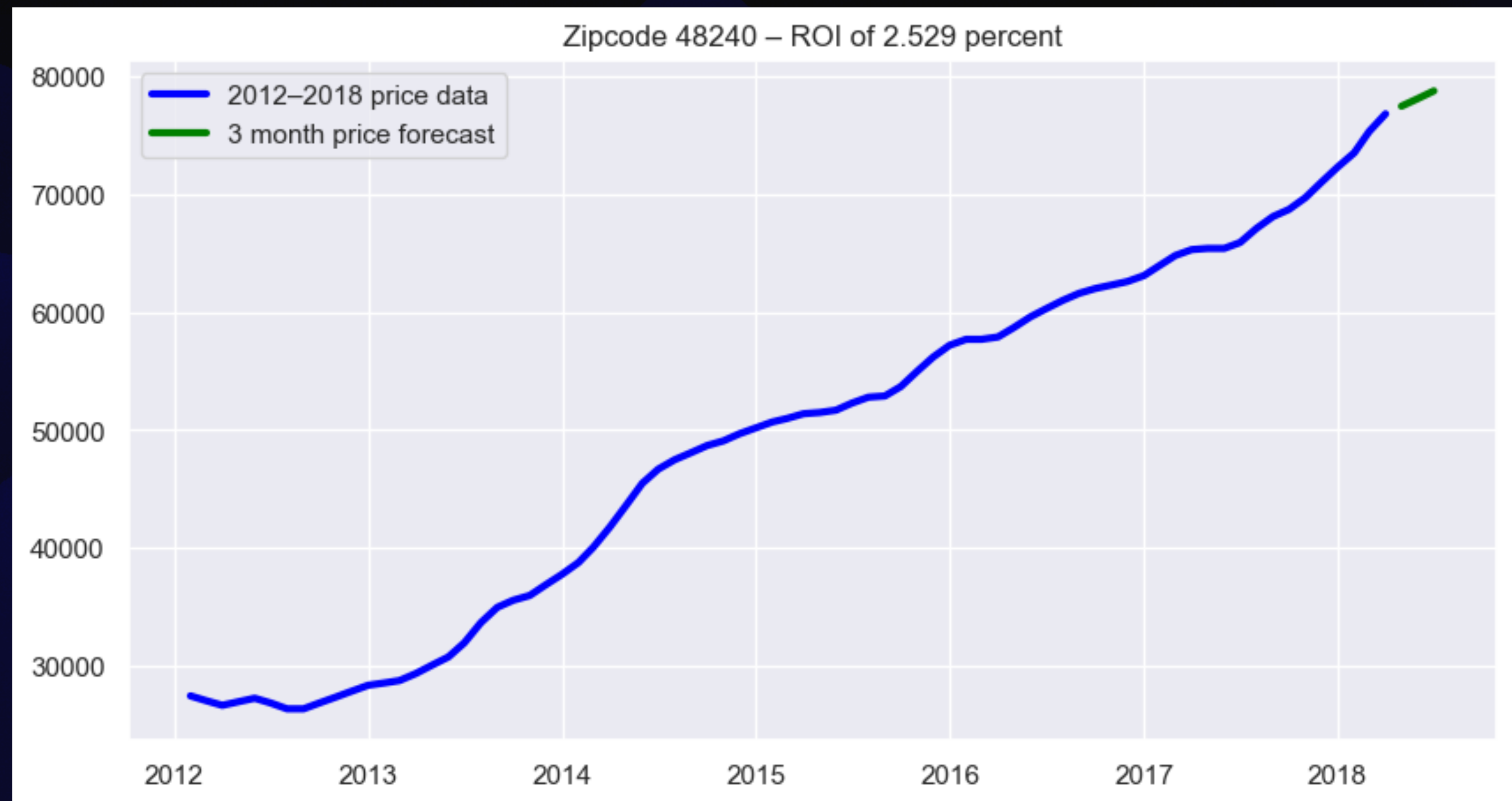Zipcode 48240 – ROI of 2.529 percent

— 2012–2018 price data
— 3 month price forecast
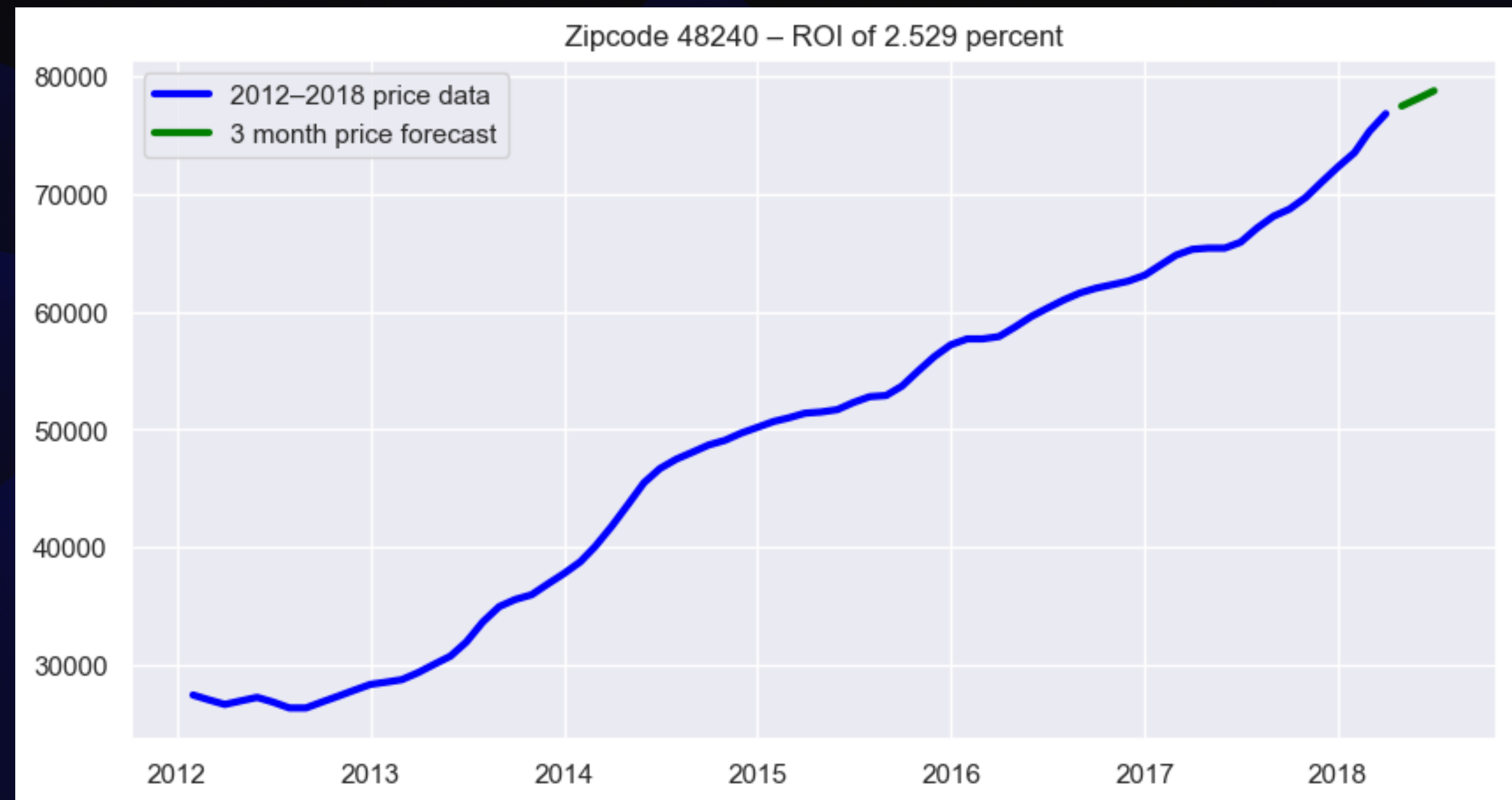
Recommendation #3 – Zip code 48237

- Location: Oak Park, Oakland MI

- Expected ROI: 2.388%

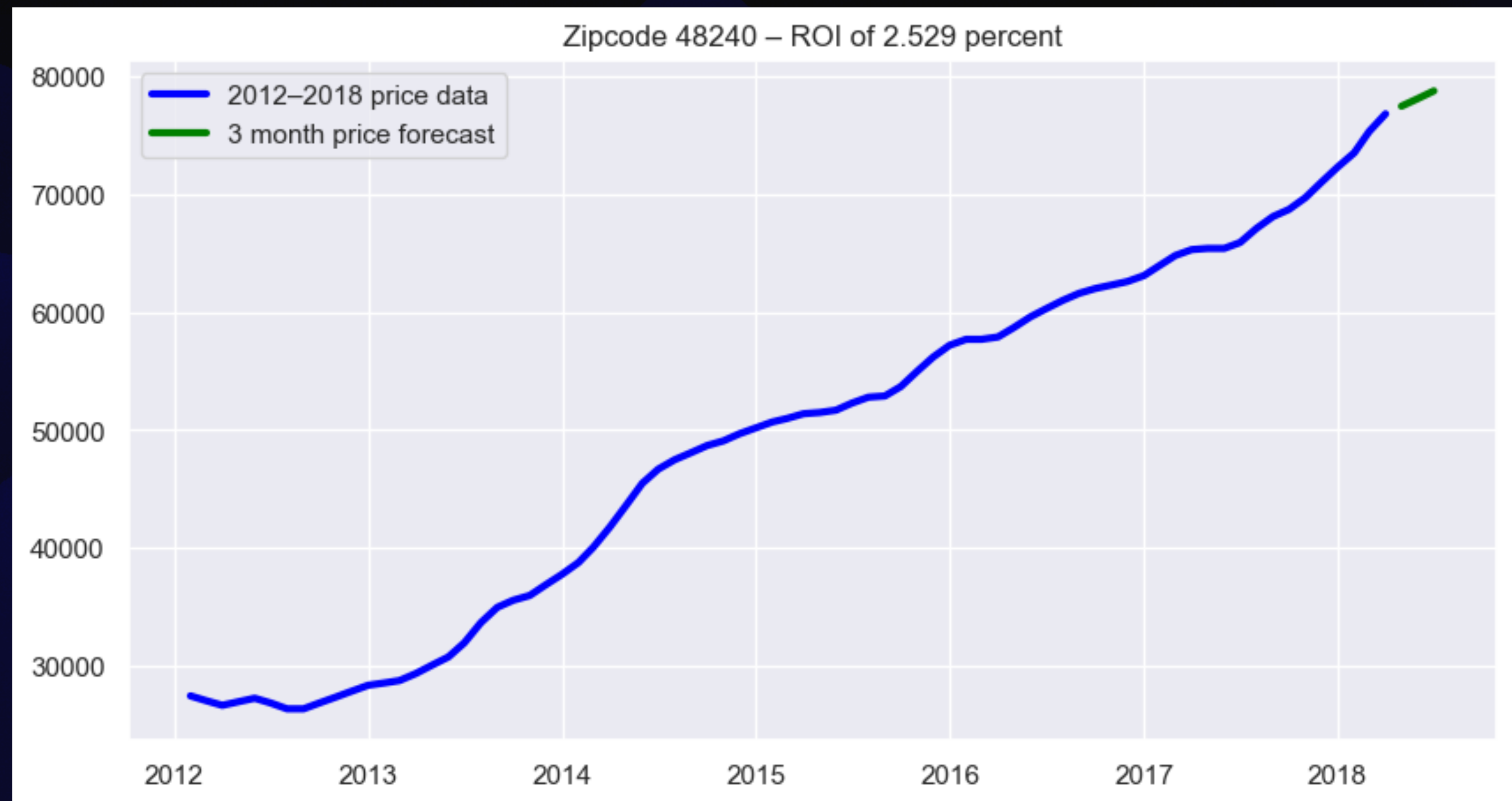- Final price prediction: $132,695

Zipcode 48240 – ROI of 2.529 percent

Recommendation #4 – Zip code 48033

- Location: Southfield, Oakland, MI

- Expected ROI: 2.377%

- Final price prediction: $144,761

Zipcode 48240 – ROI of 2.529 percent

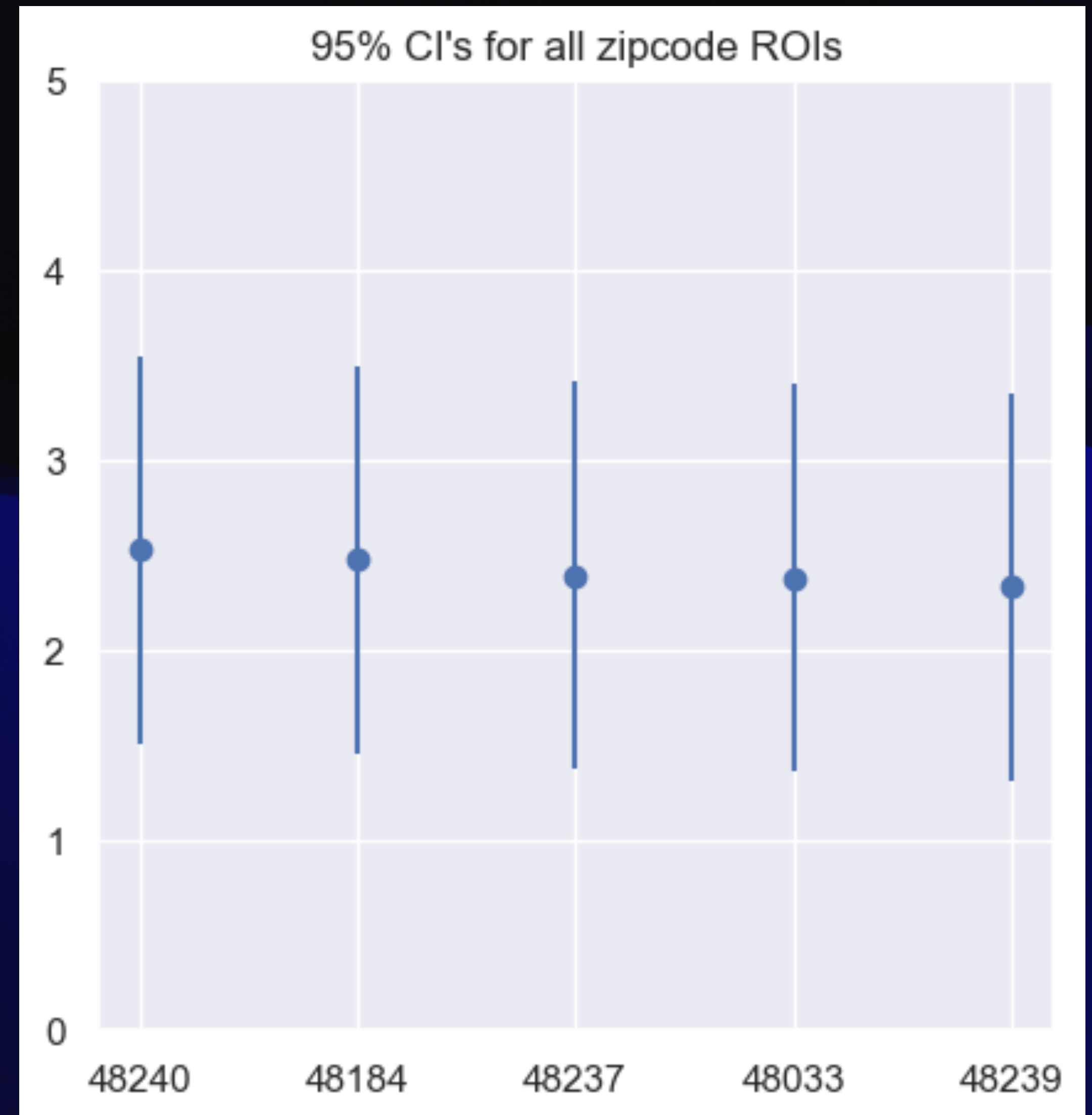# Recommendation #5 – Zip code 48239

- Location: Redford, Wayne County MI

- Expected ROI: 2.33%

- Final price prediction: $99,873

Zipcode 48240 – ROI of 2.529 percent

- 2012–2018 price data
- 3 month price forecast

Error bars for ROIs

- X-axis: zip code

- Y-axis: ROI

- ROI +/- 1%

Summary

- Five zipocdes with low volatility and between 2-3% expected ROI per month, give or take 1%

zipcode	state	county	city	average_price	volatility	ROI	price_before_investing	price_after_investing
48240	MI	Wayne	Redford	78767.92	2.351055	2.529	76800.0	78741.996645
48184	MI	Wayne	Wayne	82427.55	1.935624	2.472	89000.0	91199.835953
48237	MI	Oakland	Oak Park	101081.51	2.371316	2.388	129600.0	132695.201452
48033	MI	Oakland	Southfield	120122.26	4.695056	2.377	141400.0	144761.735864
48239	MI	Wayne	Redford	92205.28	1.521293	2.33	97600.0	99873.927678

Thank you!

- Questions are welcome.

- Contact me here:

  - angelo.turri@gmail.com