



# ZILLOW TIME SERIES ANALYSIS

Collaborators: Vivian Mosomi  
Thenya Thuita  
Mary Mwangi  
Mark Njagi

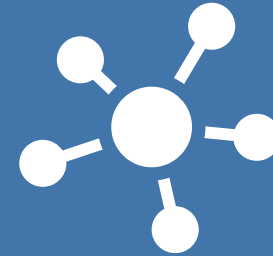
# Overview



---

Business  
Understanding

Data  
Preprocessing



---

Modelling  
& Tuning

Conclusions &  
recommendation



# BUSINESS UNDERSTANDING

## Problem statement

- Real estate involves land and buildings. Real estate investment is about strategically investing capital to generate income or increase value.
- Using data from Zillow Research, we're forecasting future property prices to assist Grand Will Realty Solutions in identifying lucrative zip codes for investment.
- As consultants, our goal is to provide valuable insights on real estate opportunities and market dynamics, empowering stakeholders to make informed decisions.



# Objectives

## **Main Objective**

To develop a time series model predicting the top five zip codes for real estate investment.

To evaluate which county exhibits the most promising real estate investment opportunities.

To understand the trend for the 5 best Zip codes to invest in based on Return On Investment(ROI)

To forecast property values over the short and long term

# Metrics of Success

- Return On Investment (ROI)
- Root Mean Squared Error (RMSE)





# DATA PREPROCESSING



## Data understanding

This data records monthly housing sales prices for zip codes from April 1996 to April 2018, sourced from Zillow.



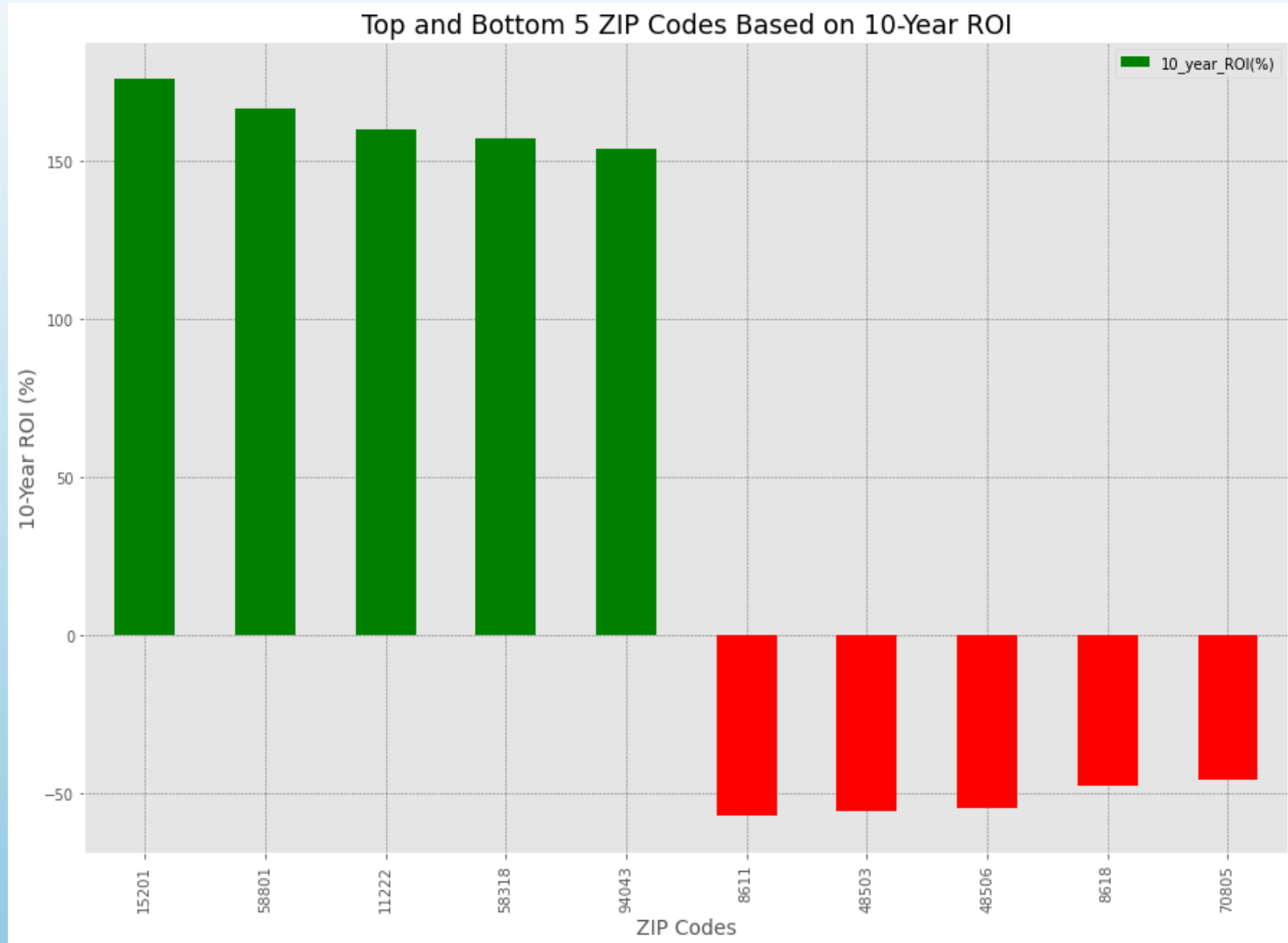
## description

- RegionID: Unique index
- RegionName: Unique Zip Code
- City: Zip code's city
- State: Zip code's state
- Metro: Metropolitan Area
- CountyName: Zip code's county
- SizeRank: Rank of zip code's size
- 1996-04 to 2018-04: Median housing sales values from April 1996 to April 2018, 265 monthly data points.
- Value: Median housing price.

# EXPLORATORY DATA ANALYSIS

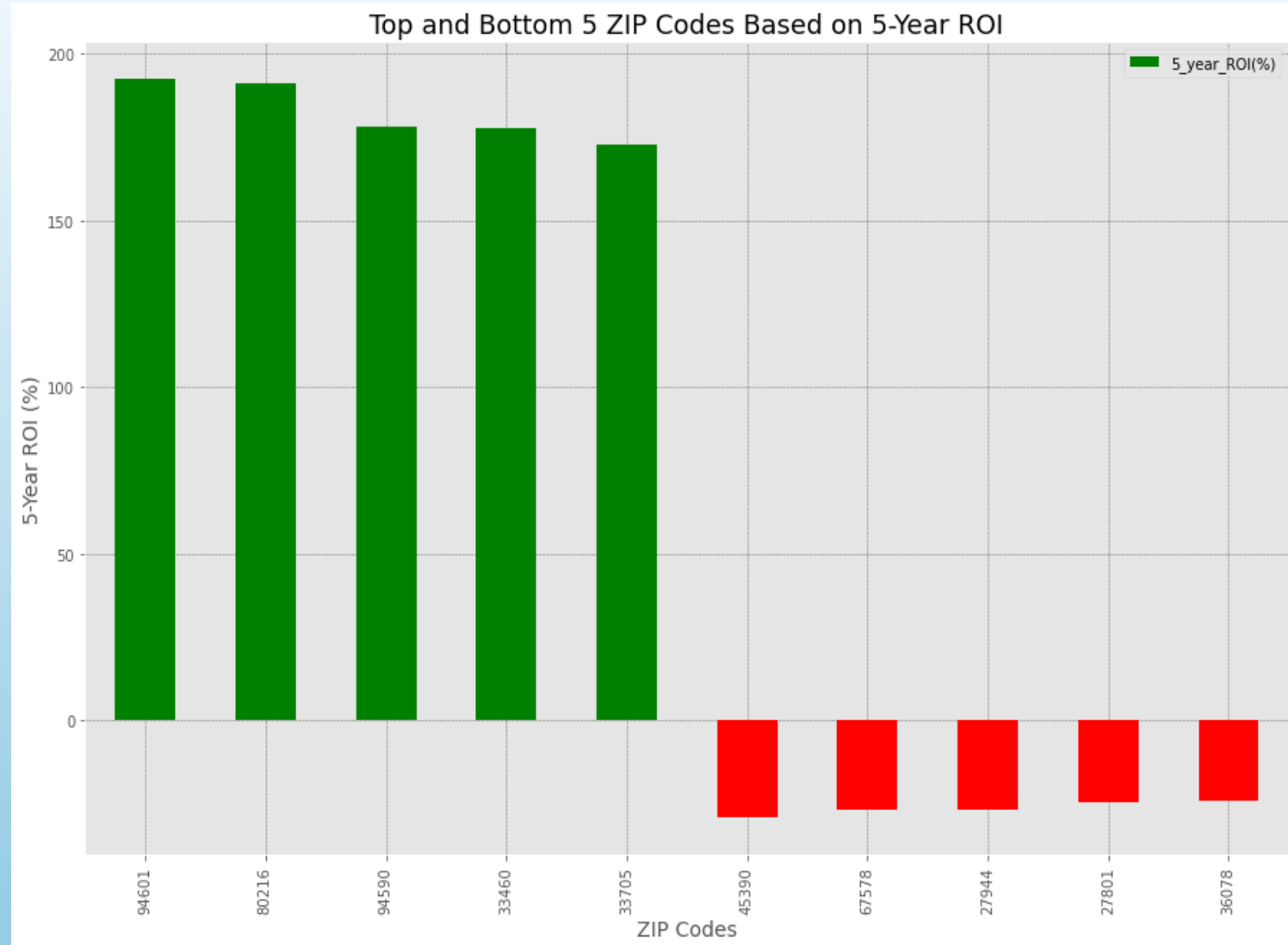
## CALCULATING THE RETURN OF INVESTMENT

Plotting the top 5 and bottom 5 zip codes for the last 10 years

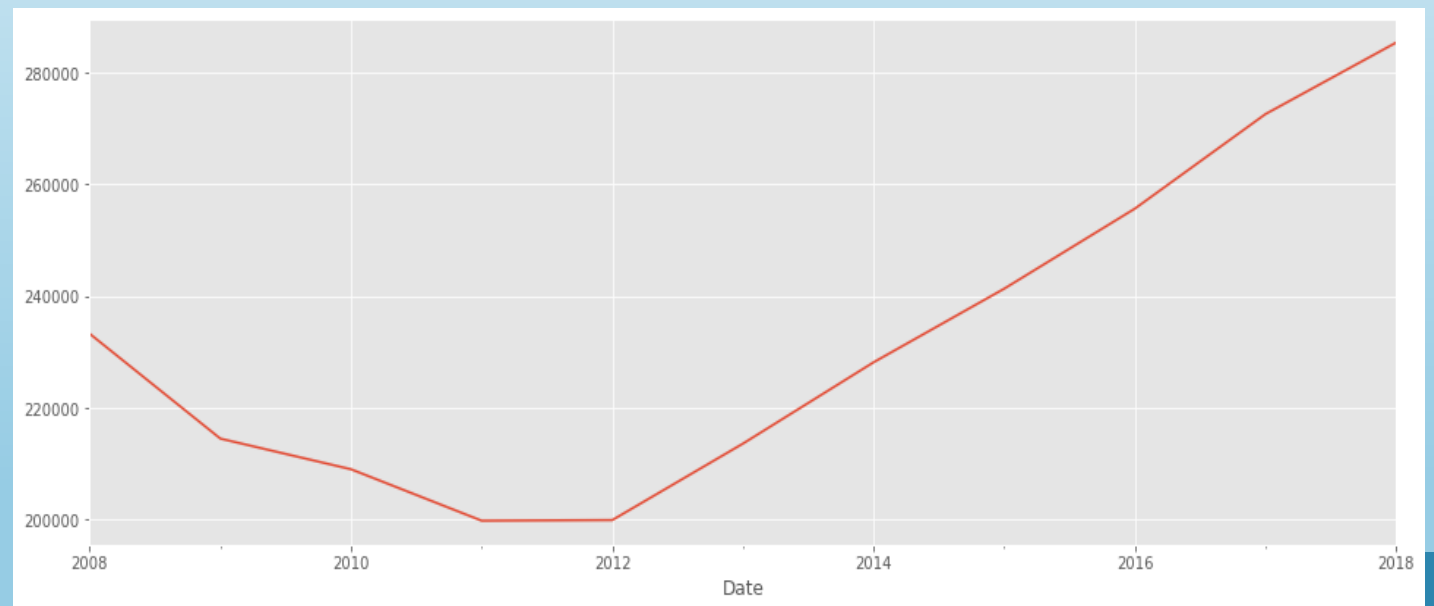
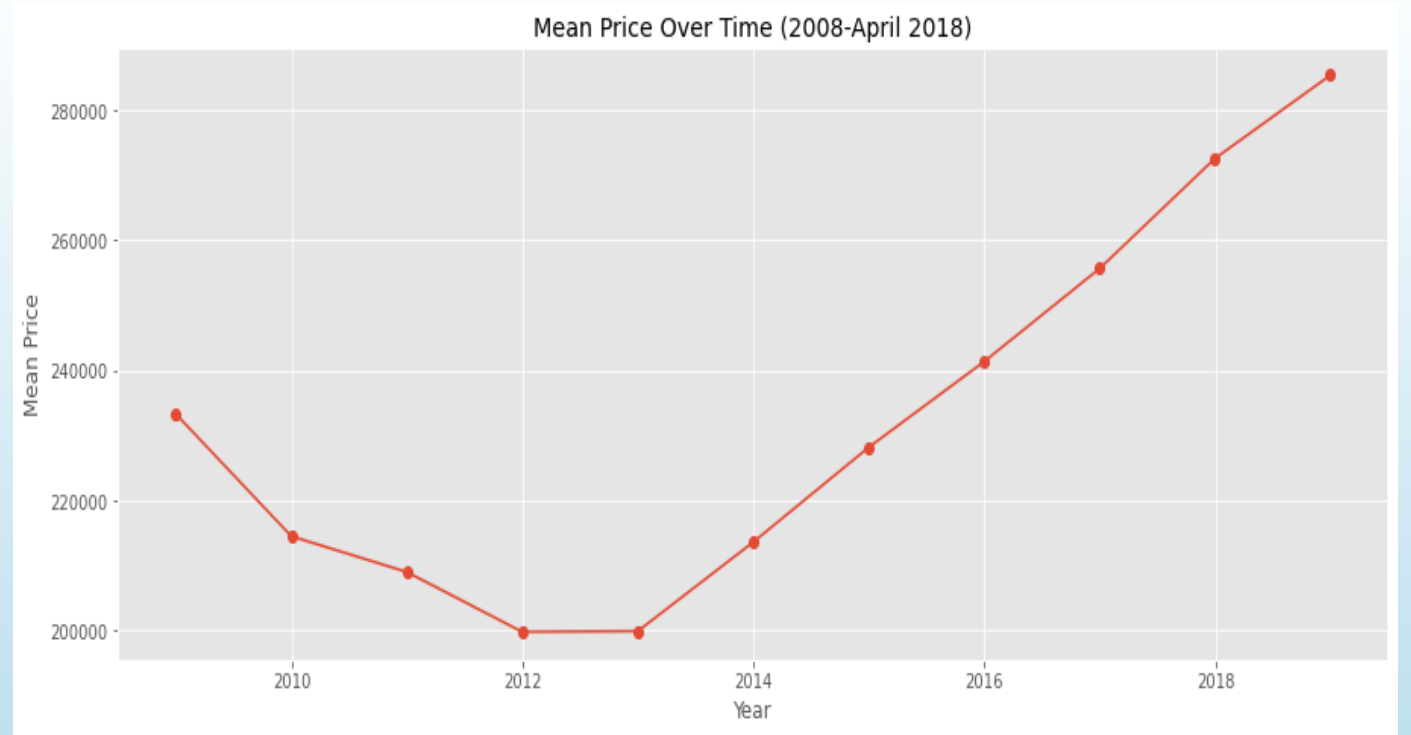




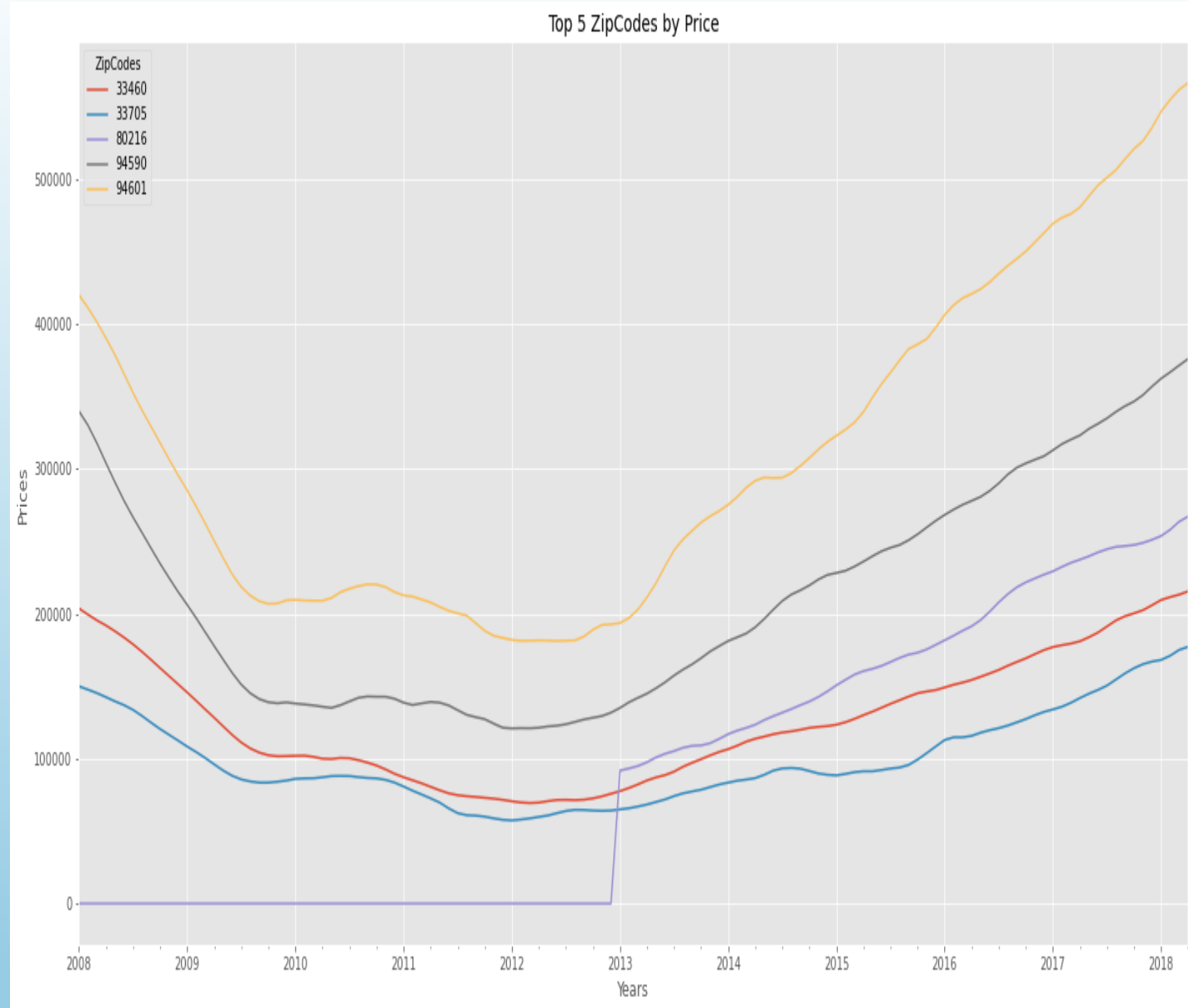
Plotting the top 5 and bottom 5 zip codes for the last 5 years



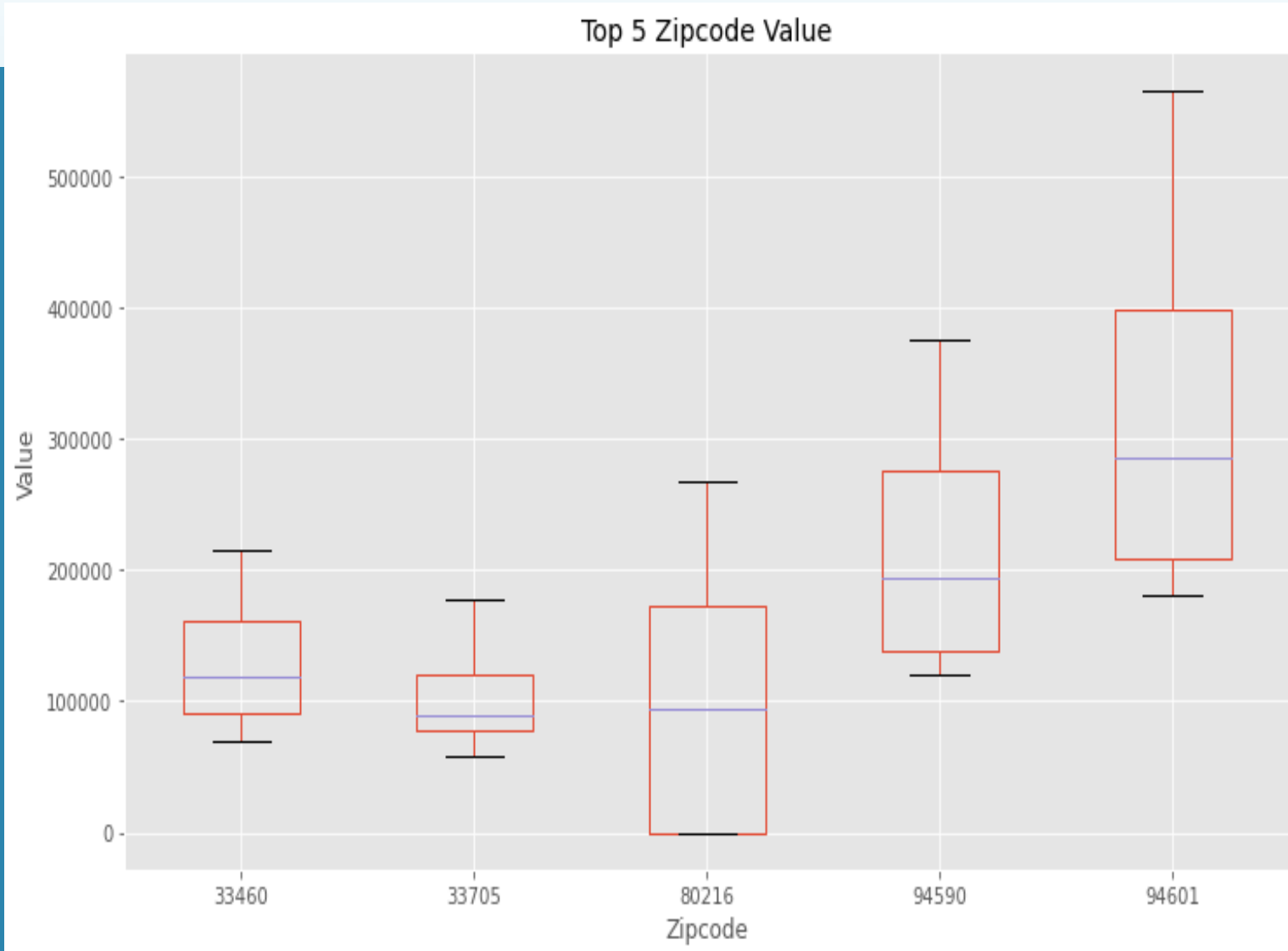
Plotting a visual to display the mean price over time



## Plotting Top 5 Zip Codes by Price



## Generating a Box Plot for top 5 zip codes



# Checking for trends, seasonality and residuals using seasonal decompose

Zipcode = 94601



Zip code =94590

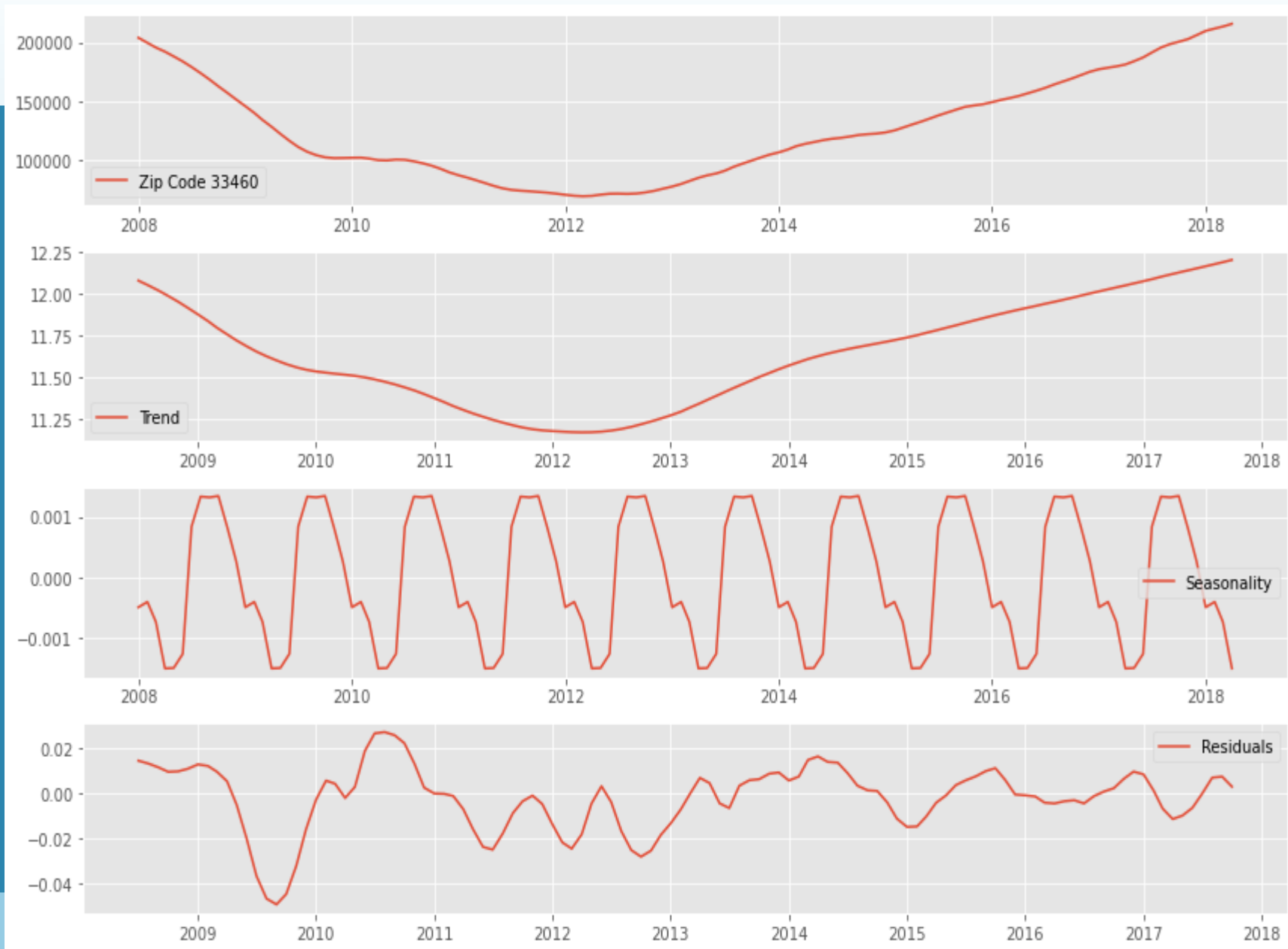


**Zipcode = 33705**

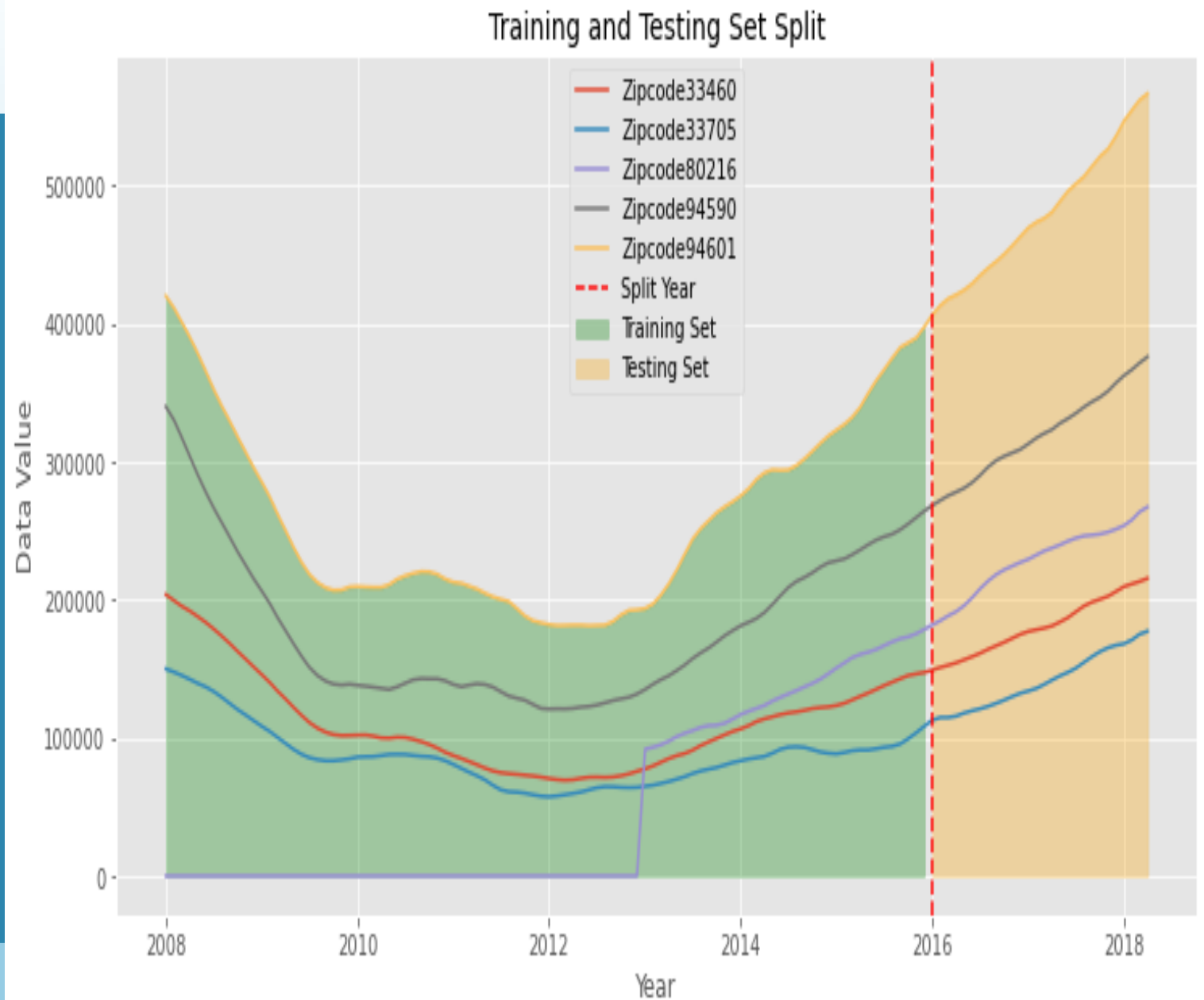




**Zipcode = 33460**



## Plot of the train-test set split



# MODELLING

ARIMA  
(AUTOREGRESSIVE  
INTERGRATED  
MOVING AVERAGE)

RMSE: 6198

FACEBOOK  
PROPHET

RMSE: 4545

SARIMAX  
(SEASONAL  
AUTOREGRESSIVE  
INTERGRATED  
MOVING  
AVERAGE)

RMSE: 1729

# CONCLUSION

-The study pinpointed the top 5 Zip Codes with the best Return on Investment (ROI) in real estate over 5 years. The real estate market trended positively, without any noticeable seasonal patterns.

The chosen top 5 Zip Codes for investment are:

- 94590
- 94601
- 33705
- 33460

-Of these, 94601 exhibits steady growth, while 33705 promises the highest predicted ROI.

Best Zip Codes with their Counties:

- 94601 (Solano County, CA)
- 94590 (Alameda County, CA)
- 80216 (Denver County, CO)
- 33705 (Pinellas County, FL)
- 33460 (Palm Beach County, FL)



# Next steps:

- Implementation of Predictive Model

- Detailed Due Diligence



- Diversification Strategies

- Continuous Monitoring



**THANK YOU**