Eddie-254 /
Phase-4-Project

`<> Code`   Issues   Pull requests   Actions   Projects   Wiki   Security   Insights   Settings

Top 5 best zip codes to invest using time series   1/5

⚖️ GPL-3.0 license

☆ **0** stars   **1** fork   👁 **2** watching   Activity

🌐 Public repository

---

main ⌄   •••

Branches   Tags

Your main branch isn't protected
Protect this branch from force pushing or deletion, or require status checks before merging. Learn more

Protect this branch   Dismiss

Eddie-254 •••   10 hours ago 🕐

View code

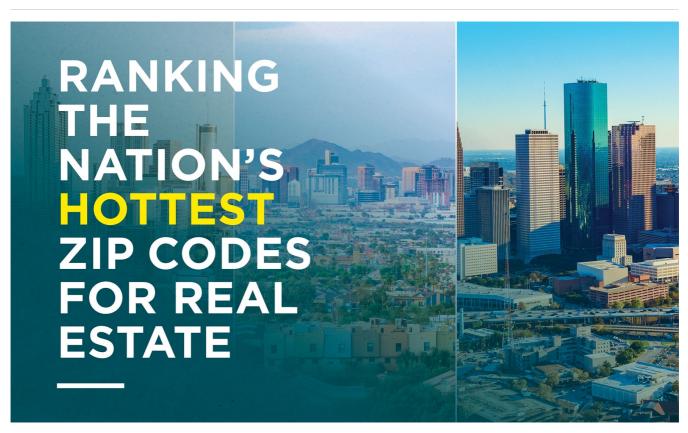☰ **README.md**   ✎

Phase-5-Project: Zillow Dataset Time Series

Contributors: Paul Kamau, Hersi Yussuf, Edwin Nderitu, Sharon Kimani, Agape Gichuki, Beatrice Kirui

Table of contents

1. Business Understanding

Overview

Real Estate Investment Firms provide comprehensive investment advisory services, including market research, property analysis, due diligence, financial modeling, and portfolio management. Our goal is to optimize investment decisions, mitigate risks, and ensure long-term success.

The primary focus of this project is to identify opportunities in real estate markets and capitalize on them to generate significant profits. We will carefully assess and mitigate risks associated with each investment based on some factors like market volatility. We will conduct a market analysis to identify areas of high demand and growth for optimal investment and prioritize investments with potential substantial returns based on factors like property appreciation and market demand.

With long-term value, our investment strategies will focus on the ability to generate consistent cashflows overtime. Real Estate Firms can achieve long-term partnerships with clients by achieving their financial objectives through successful real estate investments.

Problem Statement

At Matawi Real Estate Investment firm we seek to identify the top five areas for potential investment opportunities. The firm aims to maximize return on investment by strategically selecting areas that exhibit strong growth potential and promising real estate market conditions. By leveraging data from Zillow Research, our firm intends to make data-driven investment decisions and optimize investment portfolio.

The investment firm needs to determine the top five areas that present the best investment opportunities based on real estate market trends and historical data. We will conduct a comprehensive analysis of various factors, such as past price trends, growth rates, market demand, and other relevant indicators to identify the areas with the highest potential for future price appreciation.

Objectives

Main objective:

- The main objective is to develop a forecasting model that can accurately predict real estate price movements in different areas and assist in identifying the most favorable locations for investment between the period of April 1996 to April 2018.

Specific objectives:

- To assess and mitigate potential risks associated with market volatility and economic fluctuations.
- To Utilize time series analysis techniques to identify underlying patterns, trends, and seasonality in the real estate price data
- To Build a time series predictive model that can forecast real estate prices.
- To Evaluate the forecasting model's performance by comparing its predictions against actual real estate prices
- To forecast house prices in the next subsequent years.

2. Data Understanding

Overview

The dataset used in this project consists of historic median house prices from various regions in the USA. It covers a time period of 22 years, specifically from April 1996 to April 2018. The dataset was obtained from the Zillow website. Zillow website

Data Description

This data contains 272 columns and 14,723 rows. It has the following columns and descriptions;

`RegionID` : A unique identifier for each region.

`RegionName` : The names of the regions, represented by zip codes.

`City` : The corresponding city names for each region.

`State` : The names of the states where the regions are located.

`Metro` : The names of the metropolitan areas associated with the regions.

`County Name` : The names of the counties where the regions are situated.

`Size Rank` : The ranking of the zip codes based on urbanization.

3. Data Preparation

Within our data preparation phase, we performed the following tasks:

- Clean Data
- Checking Duplicates
- Filled Missing Values
- Feature Engineering

4. EDA

The following analysis was performed on the data:

- Univariate Analysis
- Bivariate Analysis
- Multi-Variate Analysis
- Time Series Analysis

5. Modeling

Time series modeling was chosen as it refers to building a machine learning model that can auto-generate future predictions based on existing or historical data which was our dataset. The models include;

- Auto Regressive model(Base Model)
- ARIMA model
- SARIMA model

5. Evaluation

Our success metrics are; Root Mean Squared Error, AIC and BIC. The ARMA model has the lowest AIC, BIC and RMSE hence we will use this model for forecasting.

6. Conclusions

- The housing market in New York, Los Angeles, and other top cities, states, metros, and counties demonstrates high activity and asignificant number of houses.

- Return on investment (ROI) and house prices exhibit positively skewed distributions, with outliers indicating higher returns and extremely high-priced houses.

- There is a strong positive relationship between ROI and the coefficient of variation (CV), implying that higher risk is associated with higher returns.

- New York consistently stands out as the location with the highest ROI across different levels of analysis, such as states, metros, and counties.

- San Francisco, Los Angeles, and other specific cities, states, metros, and counties offer a combination of high ROI and lower risk, making them attractive for real estate investment.

7. Recommendations

- Consider investing in real estate in New York, particularly in cities like San Francisco and Los Angeles, which have shown high ROI and relatively lower risk.

- Focus on short-term investment opportunities to capitalize on the predicted high increment in house prices for the next year.

- Adopt a conservative approach during the anticipated high decrement period of the following three years and carefully assess market conditions before making major investments.

- Consider selling properties at their peak value before the reduced increment period to maximize profits and mitigate potential losses.

- Diversify the portfolio across different locations and property types to spread risk and minimize exposure to market downturns.

- Monitor the real estate market closely, staying updated on trends, economic indicators, and regulatory developments to make informed decisions.

- Mitigate risks through thorough due diligence, assessing property quality, and conducting proper market research.

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Contributors 6

Languages

- **Jupyter Notebook** 100.0%