

# TIME SERIES ANALYSIS OF REAL ESTATE PRICES

## *Team Members*

*Joan Nyamache*  
*Marion Achieng*  
*Eugene Marius*

# SLIDE 2: INTRODUCTION

Real estate is a significant sector of the economy, encompassing residential, commercial, and industrial properties. This project focuses on analyzing the U.S. real estate market, particularly three-bedroom homes in Texas, to predict trends and identify top investment areas based on ROI.

**2020**

Suburban Boom: Price Stabilization

**2021**

**2022**

Technology  
Integration  
Expands

# PROJECT OBJECTIVE

## Main Objective:

To develop a predictive time series model to forecast real estate values and identify the top investment areas.

## Specific Objectives

- Determine the top 5 zip codes for investment.
- Analyze ROI over 1, 3, 5, and 10-year spans.
- Use coefficients of variation to assess risk.

# DATA OVERVIEW

## ***Source and Description:***

Data from Zillow includes median monthly housing prices for three-bedroom homes across the U.S., covering 23404 rows and 281 columns.

Key features include RegionID, RegionName, City, State, Metro, and CountyName.

The dataset provides valuable historical data from January 2000 to August 2022.

## CLEANING AND PREPROCESSING:

- Addressed missing values and anomalies.
- Converted dates into a standard format.
- Normalized data for consistency.
- Categorical data were encoded to numeric formats for model compatibility.

# DATA PREPARATION

# EXPLORATORY DATA ANALYSIS (EDA)

Key Findings:

## PRICE TRENDS:

Historical trends reveal significant seasonal and cyclical patterns in real estate prices.

## OUTLIERS AND ANOMALIES

Identified and addressed unusual price points to improve model accuracy.

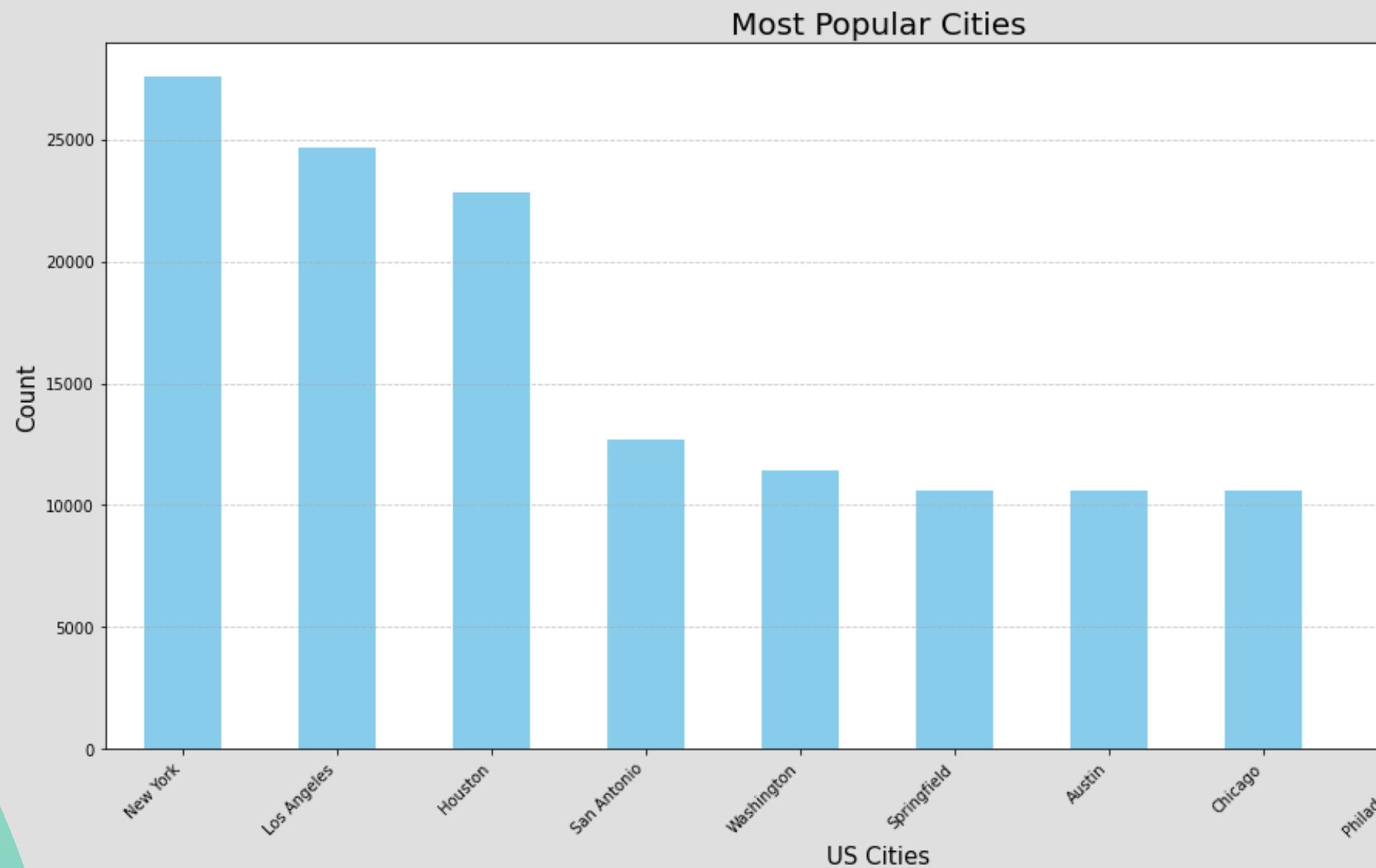
## VISUALIZATION

:

Used line plots and box plots to understand the distribution and trends in data.

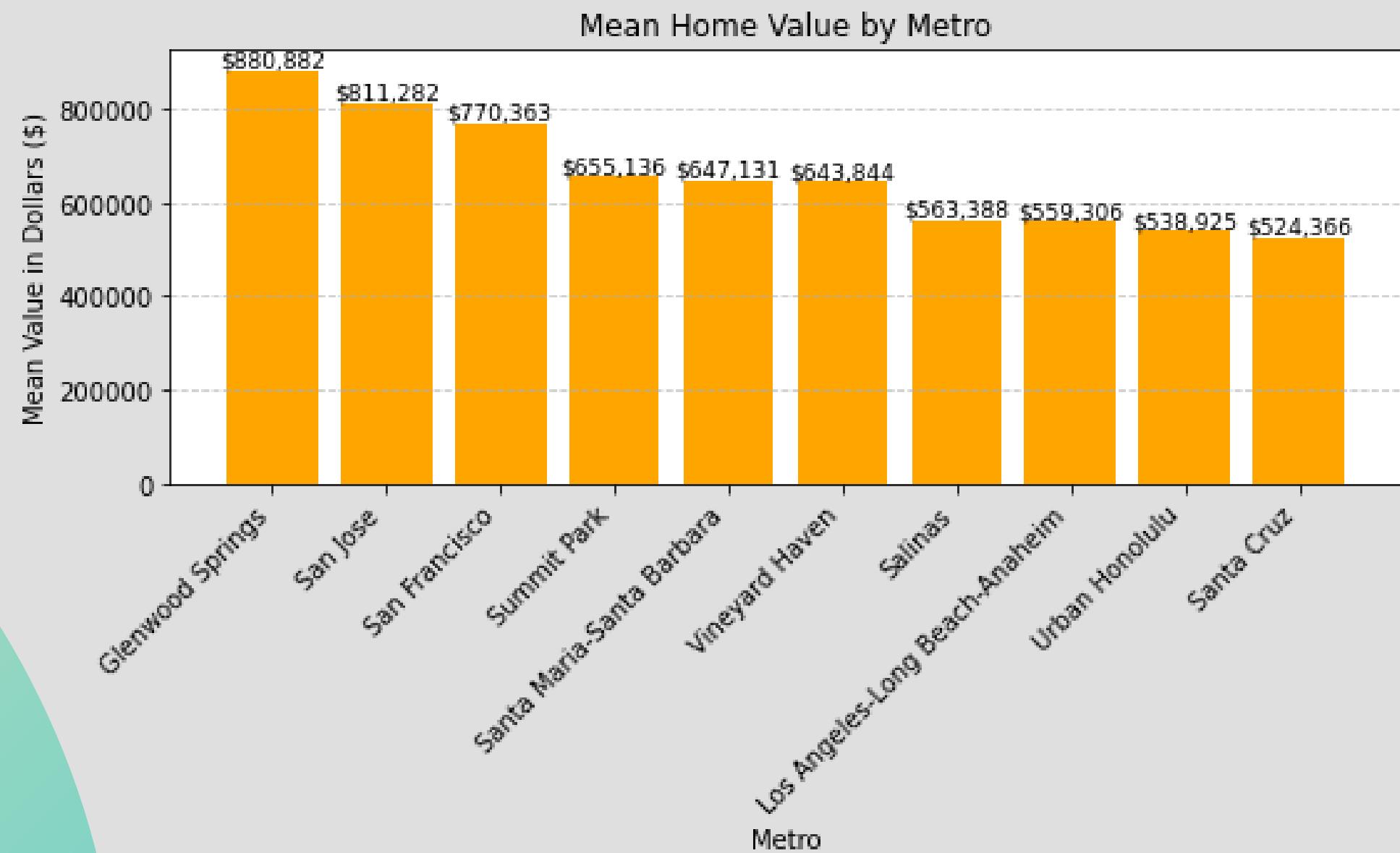
# UNIVARIATE ANALYSIS

Top 20 most frequent cities in the dataset



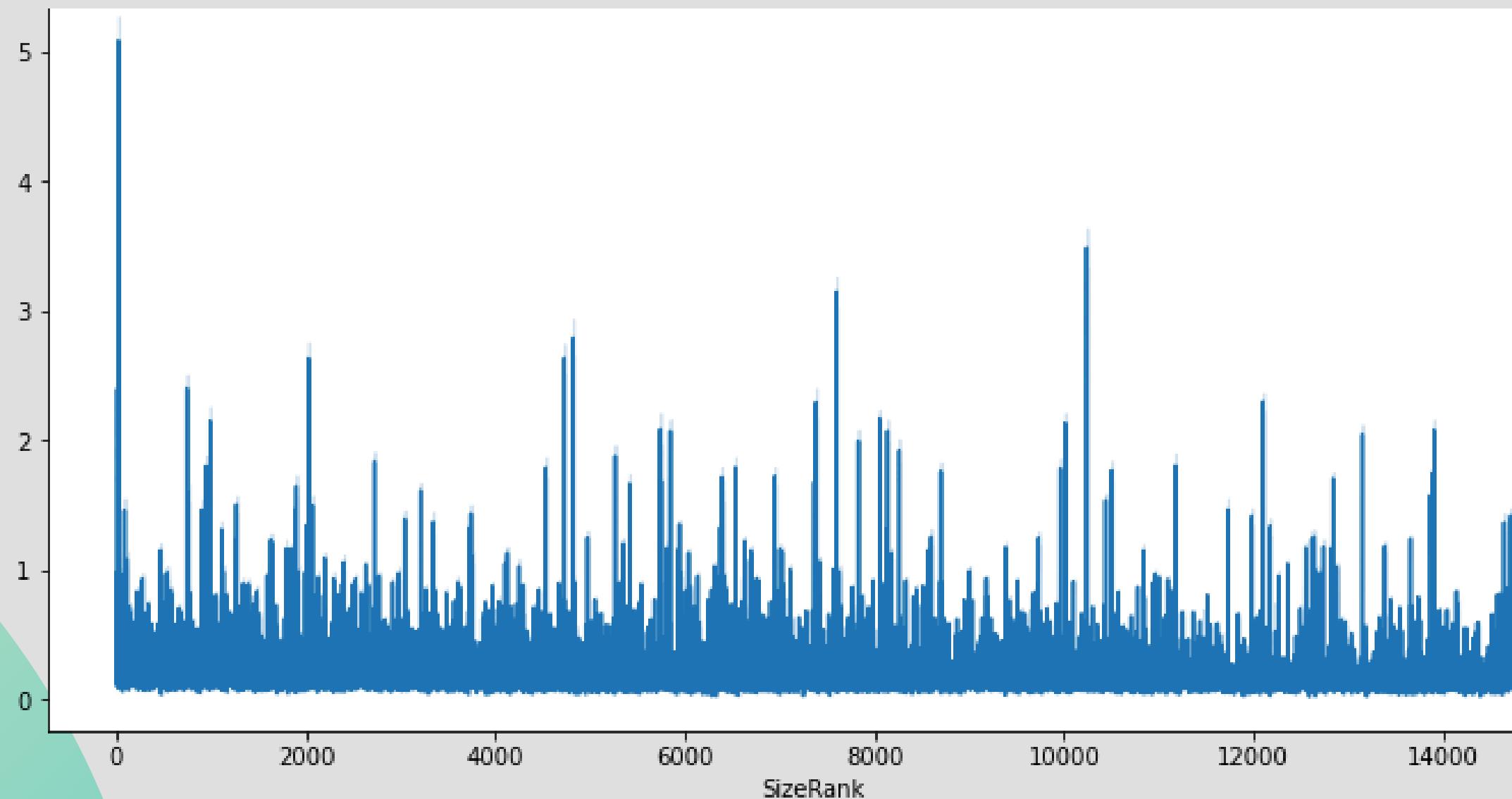
## UNIVARIATE ANALYSIS

This plot shows the top 20 most frequent cities in the dataset.

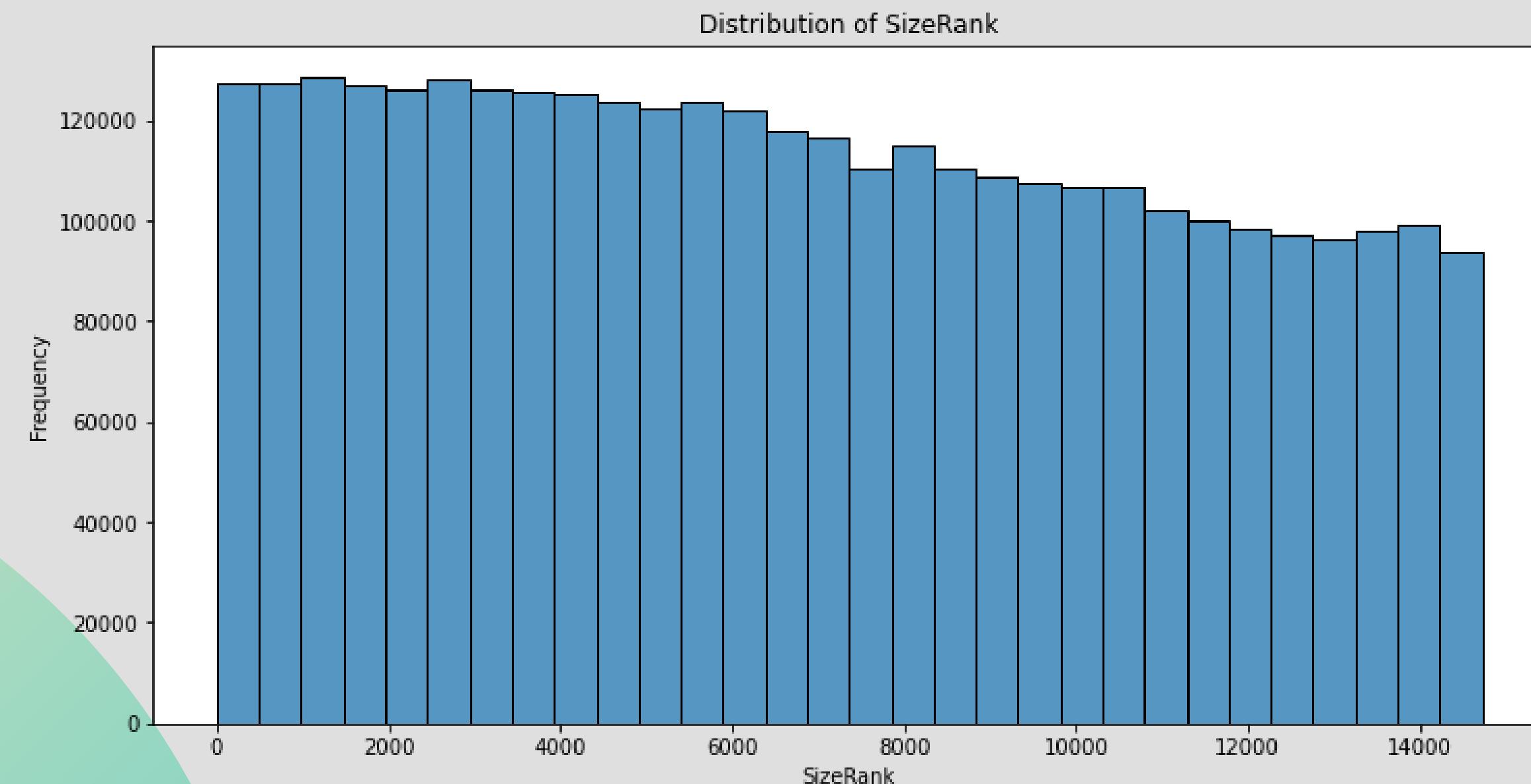


## BIVARIATE

Price vs size Rank



# DISTRIBUTION OF SIZE RANK



## MODEL SELECTION

ARIMA and SARIMA models were chosen for the effectiveness in time series forecasting.

## MODEL DEVELOPMENT

## TRAINING AND VALIDATION

Data split into training and testing sets.

Hyperparameters were tuned to optimize model performance

# PREDICTION AND INVESTMENT ANALYSIS

## **Forecasting Results:**

Predictions indicate stable growth in certain areas, with the top 5 zip codes identified based on projected ROI and price stability.

## **Investment Recommendations:**

Focus on zip codes with high ROI and low variance in price changes.  
Consider market conditions and economic factors influencing future trends.



# LIMITATIONS

## CHALLENGES:

- Data limitations, such as missing values and regional disparities.
- Model assumptions may not account for sudden market shifts or external factors.

## FUTURE DIRECTIONS:

- Incorporate more diverse data sources for enhanced accuracy.
- Explore advanced modeling techniques like machine learning for better predictive performance.

## SUMMARY OF FINDINGS:

- The analysis of the U.S. real estate market, with a focus on Texas, highlights significant investment opportunities.
- The findings emphasize the value of prioritizing high-ROI areas and diversifying investments to balance potential growth and risk.
- The predictive model developed provides a data-driven foundation for making informed investment decisions, underscoring the importance of thorough market analysis.

# CONCLUSION AND Q&A

## RECOMMENDATIONS

- Investors are encouraged to consider real estate as a viable investment option, especially in states like New York, New Jersey, Colorado, California, and Washington DC, which have shown promising returns.
- To enhance investment outcomes and mitigate risks, utilizing the predictive model to forecast future real estate values is recommended.
- This approach allows investors to anticipate market trends, make data-informed decisions, and optimize portfolio performance.

# CONCLUSION AND Q&A

# THANK YOU