

Data Assessment Document

Freddie Mac- Arizona

Akshaya Krishnan, Anirudh Murali, Rahul Murugkar, Neha Sharma, Shrushti Shah

Business Problem and Objective

Business Problem:

- The challenge at hand is creating a predictive model for Arizona's House Price Index (HPI) that can reliably detect and foresee black swan events.
- Financial crises, natural catastrophes, or any other unforeseen circumstances that have a major impact on the Arizona real estate market could be considered among these occurrences.

Objective:

- The main goal of this project is to create, put into practice, and verify a predictive model for the Arizona House Price Index, with an emphasis on spotting and reacting to black swan events.

The Data

—

Data Sources

Sources, Unavailable Data and
Data Description

The primary and secondary sources are:

- Financial Data from Federal Reserve Bank of St Louis
- Population Data from US Bureau of Labor Statistics
- Mortgage Data from Federal Housing Finance Agency

The data we could not find:

- Inflation Rate/CPI at state level

Data volume:

- 284*10
-

Data Cleaning and Preparation

Tasks performed:

- Initial dataset was from 1975-2023 with 584 rows
- We decided to focus on relevant data from 2000-2019 and filtered it.
- Performed data cleaning and cleansing of all external datasets.

Status:

- The data has been thoroughly cleaned and is now ready for analysis.



Model Consideration and Feature Engineering



Feature Selection Strategy

Strategy 1: Tree- Based Models

Tree based models like Random Forests or XGBoost can capture non-linear relationships. Their robustness and ability to handle diverse feature types, make them effective techniques for HPI prediction.

Strategy 2: Lasso Regression

Lasso automatically selects features by driving insignificant variables to 0. In the context of HPI, Lasso helps create a “sparse” model by emphasizing the most impactful features.

Strategy 3: Recursive Feature Elimination with Cross-Validation

RFECV combines feature ranking and cross validation, ensuring robust performance. Its iterative process adapts to data characteristics, making it suitable for HPI prediction where importance of features may vary.

Model Considerations

Strategy 1: Long Short-Term Memory (LSTM) Networks

Justification: Effective for capturing long-term dependencies in sequential data. Suitable for complex, non-linear relationships.

Strategy 2: SARIMA (Seasonal ARIMA) Model:

Justification: Extends ARIMA to account for seasonal patterns in the data. Effective for capturing both short-term and long-term trends.

Strategy 3: Recurrent Neural Network Model (RNN):

Justification: Captures patterns and long term dependencies in sequential data. Flexibility in modelling non- linear relationship and complex data

Status

Finished

- **Data Collection:**
Gathered data from diverse sources
- **Data Preparation:**
Completed comprehensive data cleaning
- **Exploratory Data Analysis:** Conducted EDA to uncover patterns and trends
- **Model Parameters:**
Identified 8 key parameters for model building

In Progress

- **Feature Selection:**
Ongoing process of selecting relevant features for the model.
- **Model Building:**
Actively working on building the model with the selected features to enhance model performance

Future plan

- **Model Evaluation:** We intend to use metrics such as Mean Squared Error and Root Mean Square Error to evaluate our model performance and ensure reliable predictions
- **Hyperparameter Optimization:** We plan on tuning the hyperparameter for achieving optimal performance

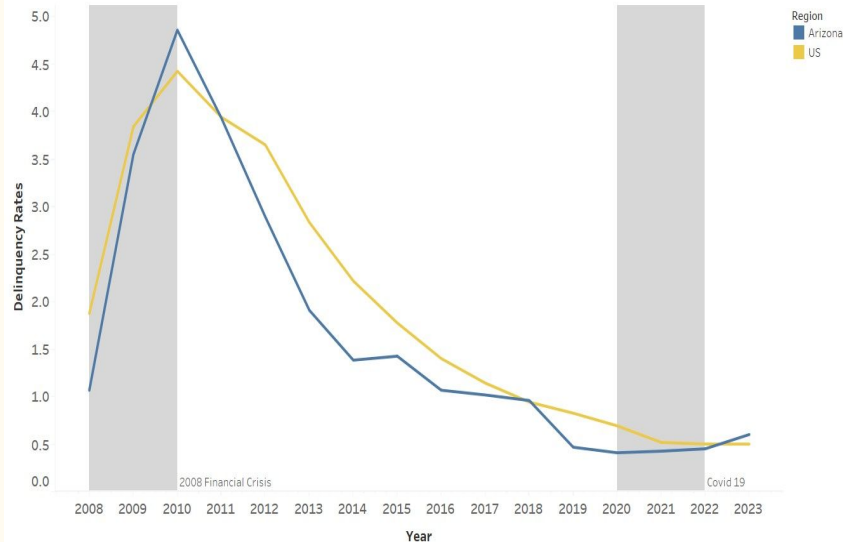
Our findings, yet.

—

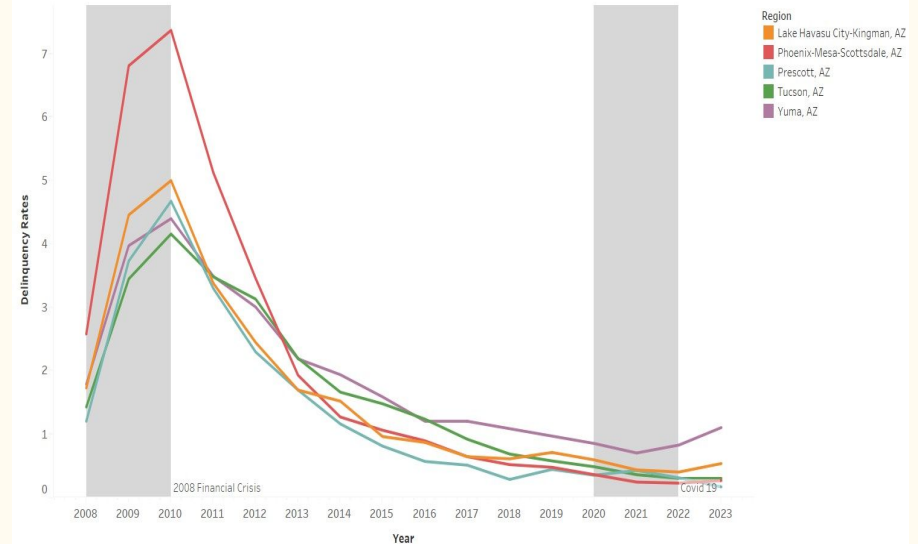
Mortgage delinquency rates

The 90+ mortgage delinquency rate is a measure of early stage delinquencies and can be an early indicator of the mortgage market's overall health. It captures borrowers that have missed one or two payments.

90+ Delinquency rate of Arizona vs the US



90+ Delinquency rates of different MSA's of Arizona



House Pricing Index

According to the National Association of Home Builders, 92.9 percent of the homes sold in the third quarter of 2021 in the Sierra Vista-Douglas metropolitan statistical area were affordable to families earning the area's median income of \$66,900.

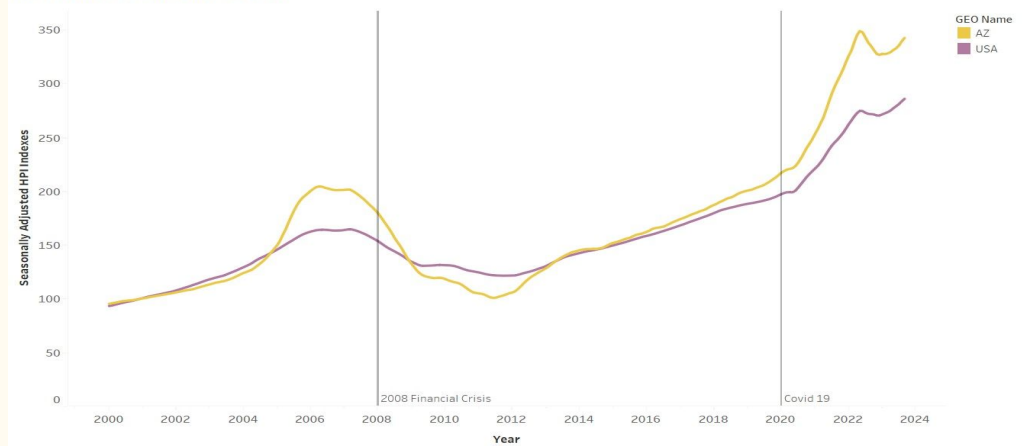
That ranked Sierra Vista area the third-most affordable housing market in the U.S., outpacing all major markets.

The index ranks the Sierra Vista area as the most affordable housing market in its region and as the only area in Arizona in the top 50 percent of the list, which includes 238 areas nationwide. The next most affordable markets in Arizona are Yuma and Tucson at 125 and 126 respectively.

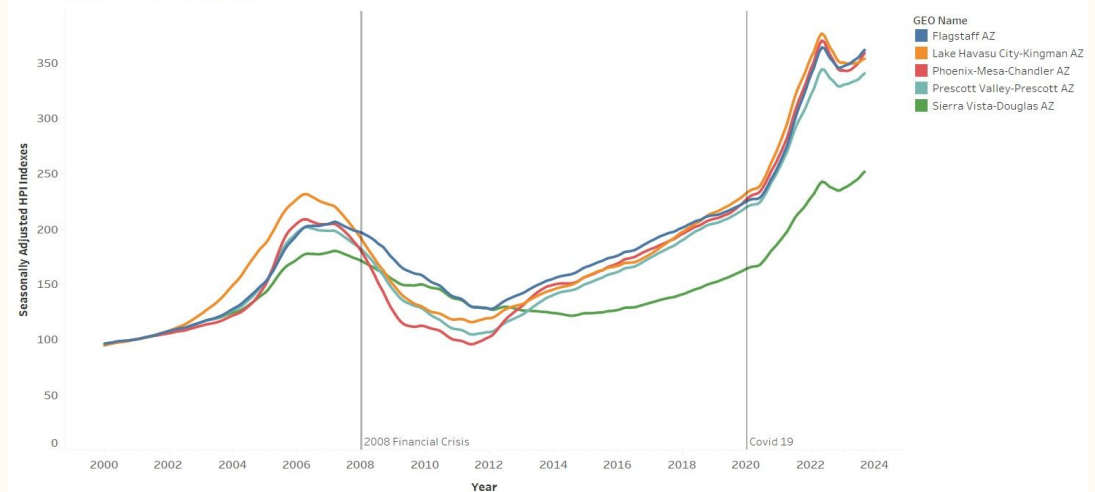
SOURCE:

<https://www.kold.com/2021/12/15/sierra-vista-area-ranked-3rd-most-affordable-housing-market/>

HPI Indexes Arizona vs US



HPI Indexes of each MSA



Unemployment Rates

Yuma, is an extreme-weather area, an agricultural area.

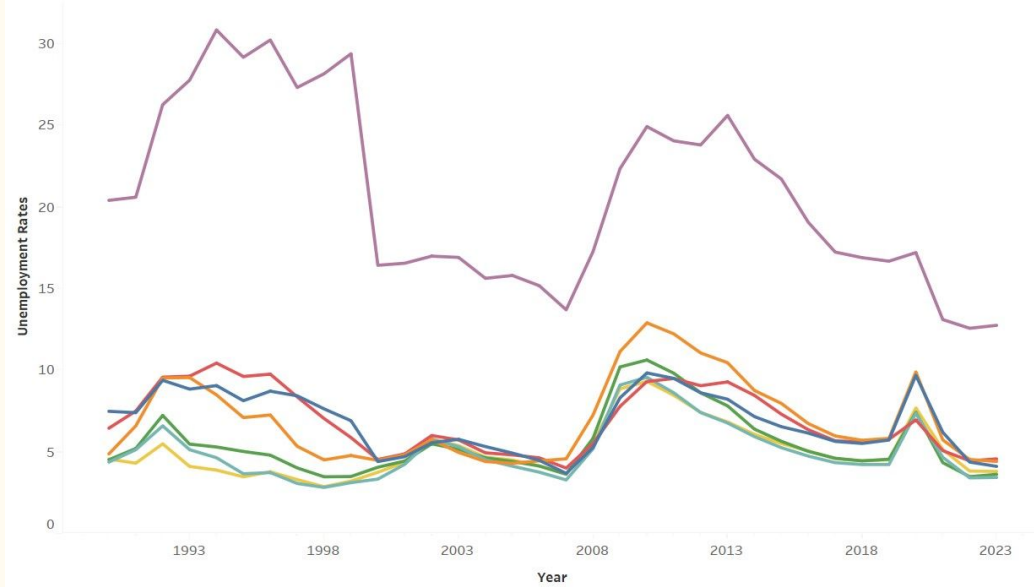
In October 2019, 153 metro areas had unemployment rates of less than 3.0 percent and 2 areas had rates of at least 10.0 percent. El Centro, California (21.2 percent), and Yuma, Arizona (16.1 percent), had the highest unemployment rates.

“We’re really tired of talking about the unemployment rate because it’s not reflective of this community,” said Julie Engel, CEO of the Greater Yuma Economic Development Corp.

SOURCE:

<https://www.azcentral.com/story/money/business/economy/2016/01/02/yuma-unemployment-agribusiness-economy/78135406/>

Unemployment Rates by each MSA



Despite high unemployment, Yuma's agribusiness continues to thrive

Will Arizona's Housing Market Crash?

Some are pessimistic about the future of the Arizona housing market, especially in Phoenix.

At the end of 2022, the National Association of Realtors predicted a 15.8% drop in combined sales and prices for the Phoenix-Mesa-Scottsdale market in 2023. Tucson was expected to fare only slightly better with a 10.2% decline.

Despite the decline in sales and available listings, home prices are on the rise. This price increase is common amongst states with lowing housing inventories. However, potential homebuyers in Arizona and elsewhere may be waiting for prices to come down and to see where mortgage rates, which have jumped considerably in the past year, go next.

What the world has to say about Arizona's housing market

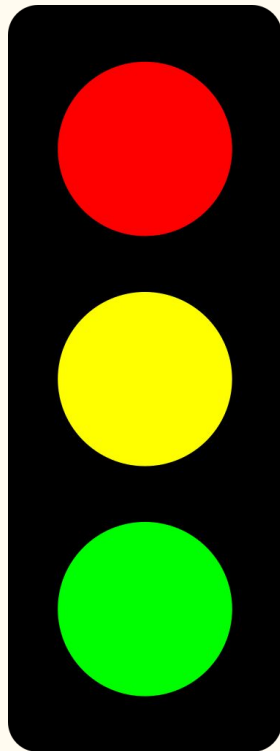
Phoenix real estate entered 2020 on a high, which should ease the pandemic's market impact

Rebuilding America: Metro Phoenix's housing market slowed in the early days of COVID-19 but is showing signs of bouncing back.

THE GAME OF SUPPLY AND DEMAND

Before the outbreak, in some Arizona communities, sellers were actively listing properties every day. Now, homeowners who may have planned to put their properties up for sale may be choosing to wait until conditions improve. With reduced supply, buyers who are still looking in these markets have fewer options. For instance, in March, Flagstaff, Arizona, reportedly saw its already low inventory of homes for sale decline. At the same time, the median home price rose by six percent.

Project Success Assessment



- Inferences and conclusions
- Feature selection
- Time series model consideration
- Data cleaning and preparation
- Exploratory Data Analysis
- Business Problem Definition

THANK
YOU