Freddie Mac - Arizona

Business Framing Document

Report by -
Neha Sharma
Anirudh Murali
Akshaya Krishnan
Rahul Murugkar
Shrushti Shah

Outline

- Introduction and Business Problem
- Background
- Data Description and Requirements
- Assumptions
- Data Limitations
- Risk and Contingencies
- Plan
- Success Criterion

Introduction and 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 Statement: 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.

Background

- Freddie Mac is a government-sponsored organization in the United States that plays a crucial role in the mortgage market by buying and securitizing loans from lenders to maintain mortgage liquidity and stability. It is essential to both enabling financing for inexpensive homes and preserving the stability of the US real estate market.
- A statistical indicator used to track variations in the price of residential real estate over time is the House Price Index (HPI). It helps with investment decisions by providing insightful information on the state and developments of the housing market. In order to compare changes in property prices and assess the success of the housing market, HPI takes into account the location, size, and type of the property.

Background (Cont.)

- Black swan events are sporadic, unpredictable, and have a big effect. The subprime mortgage crisis of 2008 and the COVID-19 pandemic of 2020 are two instances of worldwide financial crises that caused tremendous disruptions in various industries, including real estate. Other examples include seismic and tsunami catastrophes in Japan in 2011 and political developments such as the Brexit vote in 2016, both of which had unforeseen and significant effects on the real estate and financial markets.

Data Description and Requirements

- **Historical HPI Data:** Tracks past home price changes in Arizona's Metropolitan Statistical Areas (MSAs), crucial for real estate market analysis.
- **Economic Indicators:** Data from reliable sources includes unemployment rates, GDP, and income levels, providing insights into the regional economy's impact on real estate.
- **Real Estate Data:** Information on housing inventory, construction permits, and sales data from local real estate sources, aiding in supply-demand analysis.
- **Demographic Data:** U.S. Census Bureau data offers insights into population growth and age distribution, key for predicting housing demand.
- **Mortgage Data:** Delinquency rates sourced from credit agencies and government data help gauge homeowner financial stability and market health.
- **Inflation Data (CPI):** The Consumer Price Index (CPI) data from BLS assesses the impact of inflation on real estate and consumer purchasing power.

Assumptions

- The timeline we're using for building our model is from the years 1990-2022, but it is subject to change depending on certain data points which might be unavailable.
- We assume that the property is in an average condition, it doesn't account for variations based on property age, maintenance, or renovations.
- We are not taking external factors into account like political events or major infrastructure changes, rather only basing our models on historical data from previous years.

Data Limitations

- The availability of historical data at the MSA level may be limited, making it difficult to capture long-term trends and patterns accurately.
- Economic indicators, population data, and other external data sources are subject to constant updates. This can affect the consistency of the model's predictions over time.
- The availability of features, like local economic events could be limited or even unrecorded for certain time periods.

Risk and Contingencies

- **Data Integrity Risks:** Inaccurate or incomplete data
  Mitigation: Regular validation, verifying from multiple sources, and approximation methods for missing data
- **Market Risks:** Unexpected economic occurrences
  Mitigation: Constant observation and examination of the model's susceptibility to different economic scenarios
- **Model Complexity and Overfitting:** Overfitting model
  Mitigation: To optimize performance using techniques like Cross-validation, use of simpler models, and hyperparameter tuning
- **Real Estate Market Shifts:** Sudden market shifts
  Mitigation: Real-time monitoring and model flexibility for quick adjustments

Plan

1. Data Collection
Gathering historical HPI data for MSA's in Arizona.

2. Data Exploration
Analyzing the data to identify any seasonal, underlying trends, and potential outliers.

3. Feature Engineering
Identifying important features which impact house prices like GDP, inflation, population growth, interest rates, and more

4.Model Selection
Regression models like Linear Regression, Decision Trees, Random Forest, Gradient Boosting will be used

5. Training, Validation and Performance
Splitting the data based on time. Evaluating performances based on metrics like Mean Absolute Error, Root Mean Square Error

6. Adjusting for Seasonality
Making more accurate predictions by adjusting for seasonality and removing any repetitive patterns that can mislead our model

7. Monitor and Update
Regularly monitoring our model's performance and updating it as the new data becomes available

Success Criterion

- The model should predict housing prices with a high degree of accuracy. Metrics like Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) can be used to quantify this.
- The model should have the capacity to assess and predict market trends, such as the likelihood of rising or falling housing costs in a certain location. This can assist homeowners and investors make prudent decisions.
- The model must be based on high-quality and up-to-date data. Data accuracy and integrity are crucial to the success of the model.
- The model's effectiveness can be evaluated based on its coverage of diverse housing markets. A model that can provide insights and predictions for a wide range of geographic areas is more valuable.

THANK YOU