Predicting Investable Cities by Population and House Prices

Introduction:

Cities like New York City, Boston, and San Francisco have seen up to a 250% increase in house prices over a period of 30-40 years. Over that same period of time, the population in these cities have grown about 5%-25%. Every year, new residents of these cities are having to pay more for houses compared to residents from previous years, making perspective investors interested in buying homes in these major cities as a result driving prices up.

Property investors are looking to purchase investment properties; properties that are not their primary residences and are bought with the purpose of "earning a return on investment." There is not a big risk to property investors investing in properties in big cities, however their return on investment may not be as big as they would like.

What is the Problem I Want to Solve:

The first question I want answered is, What makes a city a more investable city for the longer term? Is it the population size, house prices, or a mix of both? If population and house prices are indicative of labeling a city as an investable city, then is it possible to identify a city that is currently considered a city not worth investing in, but in the future can provide great ROI?

I want to create some transparency for homeowners. Homeowners place trust in their hired real estate agents to determine the proper price of the house they are selling. But if they can get an idea for the price of their house then they wouldn't have to blindly trust real estate agents.

Who are my Clients, and Why do they Care about this Problem? :

I believe there are two clients who will benefit from this problem being answered:

1. Property Investors
    a. As I mentioned earlier, property investors can benefit the most from this because homes in major cities are limited while their prices are through the roof. Property investors are usually large corporations or individuals with a lot of money and are actively looking for investment properties. This project will help these investors choose an investment property where the ROI will be greater compared to properties bought in major cities.
2. Home Owners
    a. Home owners can also benefit from having this question answered. By taking a look at future trends, they will be able to figure out the most optimal time to sell their house, if they are looking to move.

What Data am I Using? How will I Acquire the Data :

The dataset I am using is from a previous Kaggle competition. It is called [House Prices: Advanced Regression Techniques](#). The Kaggle dataset contains house prices and house features for houses sold within the years 2006 to 2010 for the town of Ames, IA.

How will I Solve This Problem:

Both the datasets need to be cleaned for proper data wrangling. I believe this is modeled as a Regression problem and I will check to see what features determine the house price. I will be using those features in order to predict the house price of a new house that comes on the market.

What Will be My Deliverables?:

I will include my Jupyter Notebook with all my code and comments, a paper, or if time allowing a blog post.