Team: House Price
Project name: REPRICE
Members:
Kunal Matthews
Mohammad Heydary
Don Feng

Project Description:
REPRICE is a web-based Housing Price estimator based on physical and geographical attributes of a property. The system leverages machine learning to predict the estimated median value of a property at different levels. With the results, the user will be able to decide on whether to buy, sell or wait for a better market price.

Modeling Detail:

REPRICE utilizes the Zillow research dataset which is freely available to the public. Zillow's data is broken into multiple subsets each of which addresses an attribute that is thought to be significant in determining the value of a property. REPRICE will use the following subsets:

1. State time series: a list of historical median property values in a state through time
2. County time series: a list of historical median property values in a county through time
3. City time series: a list of historical median property values in a city through time
4. Zip code time series: a list of historical median property values in a zip code through time
5. Neighborhood time series: a list of historical median property values in a neighborhood through time
6. House information: property specific attributes such as the number of bedrooms and square footage

The dataset provided by Zillow contains historical data going back to 1998, however many values are missing. Thus, we shall only use the more recent data that has minimal missing values to make the analysis feasible. For modeling, we plan to break the process into multiple models each predicting at one level in a serial format (i.e., one model's output is appended to the input of the next model in the series). A unified modeling approach is not feasible due to the chaotic and complex behavior of the housing market time series at each level. The final prediction of the series of models along with previous predictions will be mapped to a human-readable format and presented to the end user. The details of modeling are subject to change as we do more experimentation on the data and discover potential changes that could reduce the overall prediction error.

The goal for REPRICE is to create a functioning software that can take the users input, preprocess it, and pass it through the developed model to generate a reasonably accurate prediction of the property price with the goal of providing a better insight to the end users.