

Team House Price REPRICE Use Case Report

Revision History

Authors	Description of Change	Sections	Rev	Date
Mohammad H., Kunal M., and Don F.	Initial Release	All	0	2/13/2019
S. Graham	Changed use case list and description to be same as diagram.	2.2	A	2/16/2019

Table of Contents

1	Team Description	4
2	Project Description	5
2.1	Use Case Diagram	7
2.2	Use Case List	8
2.2.1	Input Property Geographic Information	8
2.2.2	Input Property Specific Information	8
2.2.3	Query for most recent records	9
2.2.4	Predict Property Value	9

1 Team Description

Team Member Name	Email Address
Mohammadreza Hajy Heydary	mheydary@csu.fullerton.edu
Kunal Matthews	amos-m2345@csu.fullerton.edu
Don Feng	donfeng97@csu.fullerton.edu

2 Project Description

Statement of Purpose:

Design and implement a web-based housing price estimator based on physical and geographical attributes of a property.

Detailed 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:

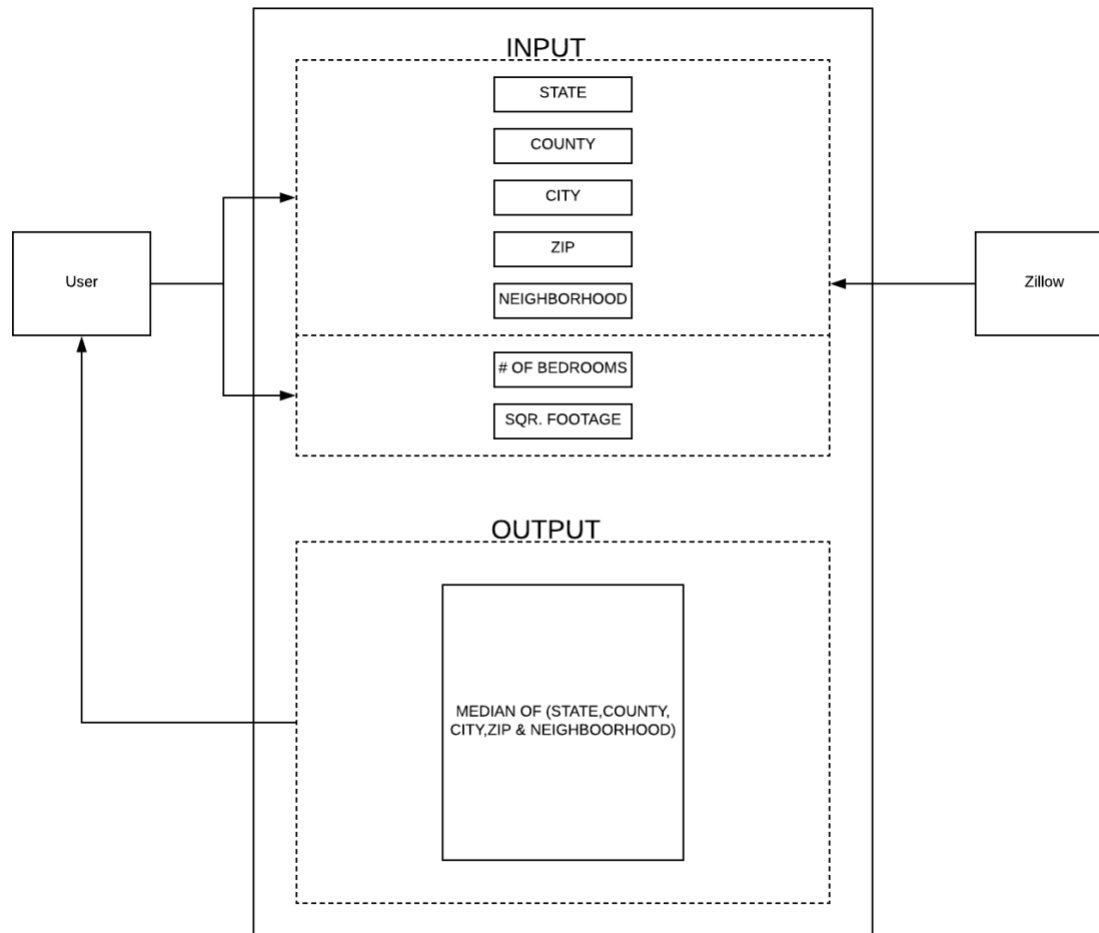
REPRICE utilizes the Zillow research dataset which is freely available to the public. Zillow's data is broken into multiple subsets each of which addresses an attribute that is thought to be significant in determining the value of a property. REPRICE will use the following subsets:

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.

2.1 Use Case Diagram



2.2 Use Case List

Use Case		
Sequence Number	Actor	Goal
1	User	Input property geographical attributes
2	User	Input property specific attributes
3	User	Predict Property Value
4	Zillow	Query for the most recent records

2.2.1 Input Property Geographic Information

Primary Actor: User

Secondary Actors(s): Zillow

Goal in Context: Input valid geographical information

Preconditions: The geographical attributes must exist in the dataset employed for development of the core model

Additional Description:

User input validation happens locally, and user cannot move to the next steps unless all the necessary requirements are met.

2.2.1.1 Error Case

If the dataset that the user inputs does not exist, then we will tell the user that it does not exist and ask them to enter another set of data.

2.2.2 Input Property Specific Information

Primary Actor: User

Secondary Actors(s): Zillow

Goal in Context: Input valid property specific information.

Preconditions: Number of bedrooms shall be provided in a categorical format ranging from one bedroom to five and more (i.e., 5 categories in total)

Additional Description:

User input validation happens locally, and user cannot move to the next steps unless all the necessary requirements are met.

2.2.2.1 Error Case

If the dataset that the user inputs does not exist, then we will tell the user that it does not exist and ask them to enter another set of data.

2.2.3 Query for most recent records

Primary Actor: Zillow

Secondary Actors(s): none

Goal in Context: update local database with most recent data

Preconditions: System has already been preloaded with Zillow historical data

Additional Description:

Periodically, Zillow database will be queried for the most recent months data. When this data is available it will be added to the database..

2.2.3.1 Error Case

If newer data is not available, no update will take place.

2.2.4 Predict Property Value

Primary Actor: User

Secondary Actors(s): Zillow

Goal in Context: Provide a point estimate of the property value based on the user input and recent trends in property values.

Preconditions: Valid data for the property to value is loaded.

Additional Description: none

2.2.4.1 Error Case

none