HTU Upskilling Program : DataScience Track

Capstone Project : Can you predict the price of a house

Student Name : Shahed Saeed Ahmad Ahmad

Date : 15th Jan 2023

<https://github.com/ShahdWard/Technical-CapStone>

* Introduction

A house price is the amount of money that a person will pay or receive in exchange for a house or property. The price of a house can vary depending on a number of factors, including the location, size, and condition of the property, as well as the state of the real estate market. Factors like the economy, interest rates and supply and demand also play a role in determining house prices. The price of a house can be influenced by many factors, including the location, condition, size, and age of the property, as well as the state of the real estate market. House prices can also be affected by changes in the economy, such as interest rates, inflation, and unemployment.

* Problem

Predicting house prices can be a challenging task because of the many factors that can influence the price of a property. Some of the common problems in predicting house prices include:

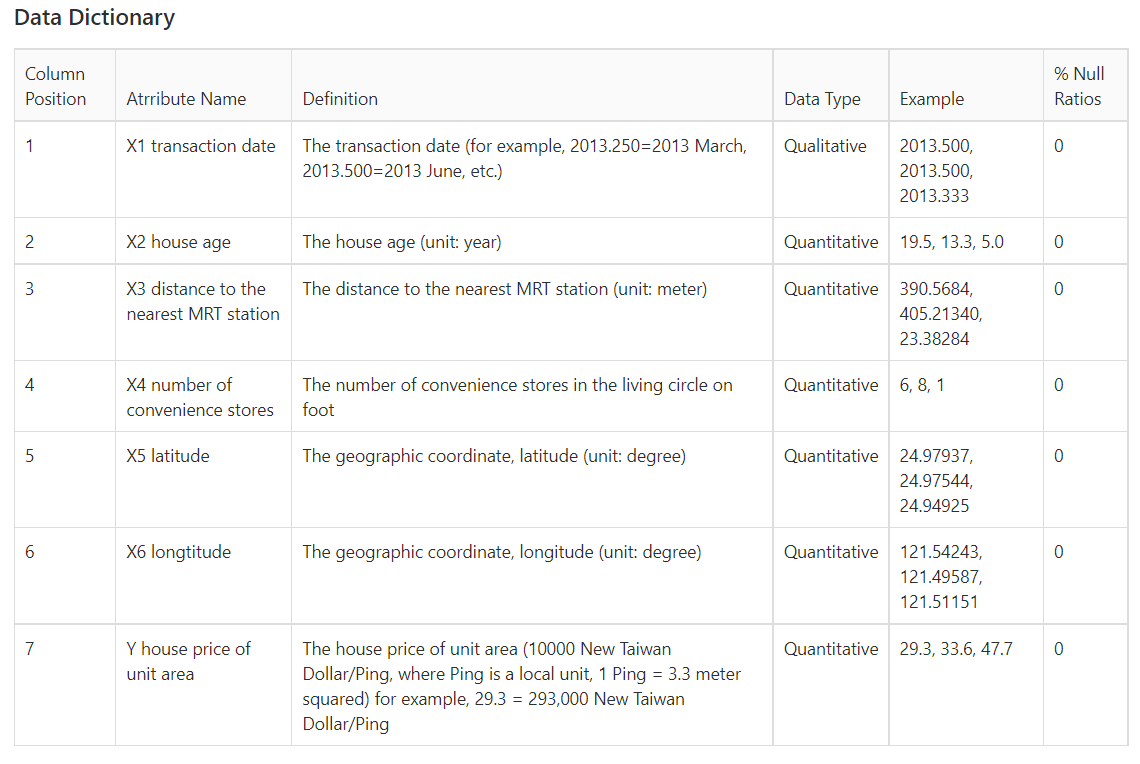
1. Lack of data: In some areas, there may be a limited amount of data available on house prices, making it difficult to create accurate predictions.
2. Heterogeneity: Houses can vary greatly in terms of size, age, condition, and other factors, making it difficult to compare them and create a single prediction.
3. Changing market conditions: The real estate market can change rapidly, making it difficult to predict future prices based on historical data.
4. Multi-collinearity: Some features are correlated with each other and this can cause problems in prediction.
5. Outliers: Some houses may have unique characteristics that make them difficult to classify, such as a large lot size or a unique architectural style.

Despite these challenges, there are various techniques and models that can be used to predict house prices, such as linear regression, decision tree, Random Forest, Gradient Boosting, etc. These models can be trained using historical data on house prices and other relevant factors, and then used to make predictions on new properties.

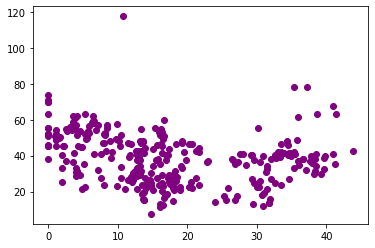
* Data set

Source: data science dojo

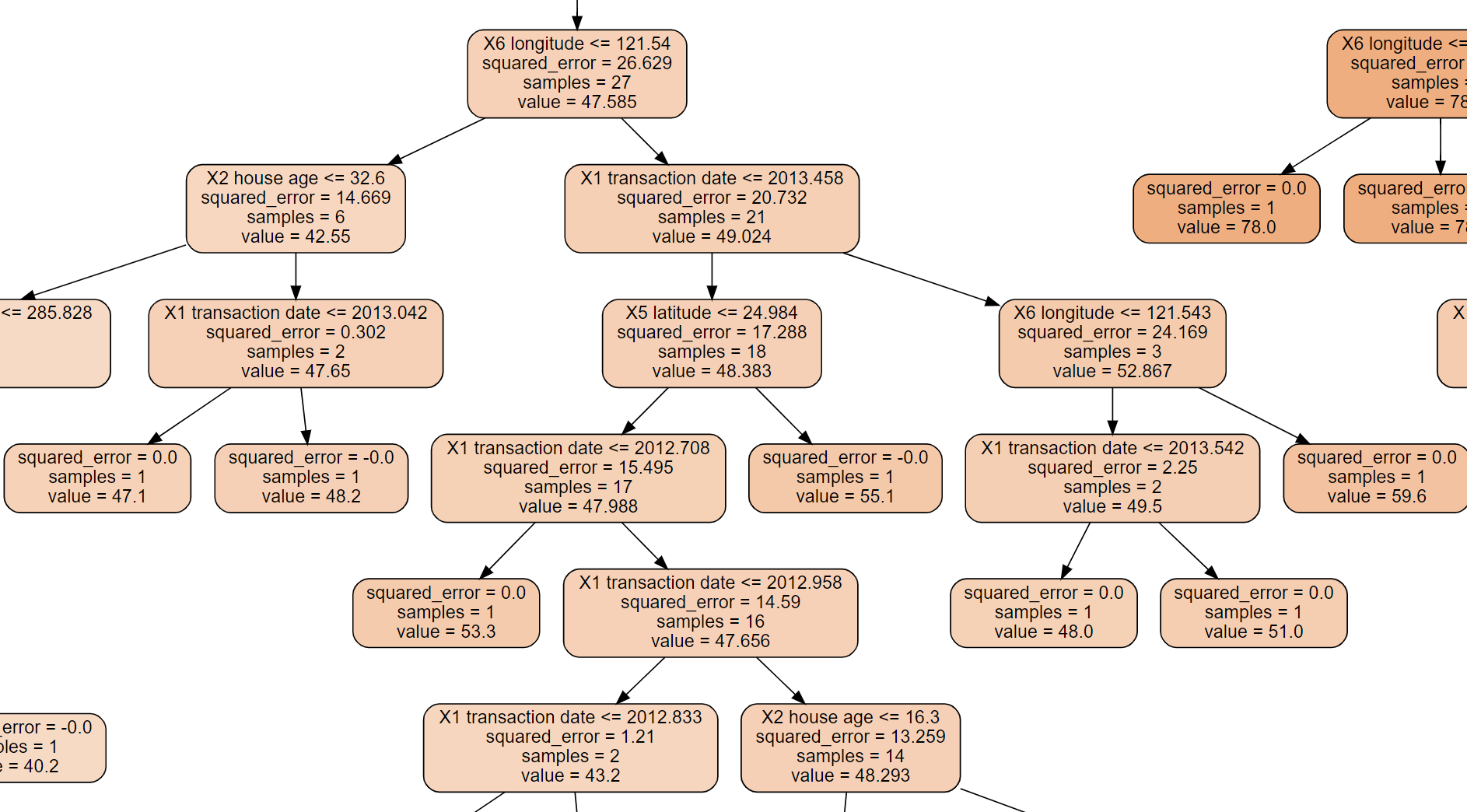
Link : <https://code.datasciencedojo.com/datasciencedojo/datasets/tree/master/Real%20Estate%20Valuation>



* Data visualization
  + Python



**Figure 1 Linear Regression**



**Figure 2 Decision Tree**

* Model

# **Linear Regression**

* + **Decision Tree**
* Results
  + 58.9% (Linear Regression model)
  + 55.5% (Decision Tree model)
* References

Yeh, I. C., & Hsu, T. K. “Building real estate valuation models with comparative approach through case-based reasoning. Applied Soft Computing, 65, 260-271, (2018).