HTU Upskilling Program : DataScience Track

Capstone Project : Can you predict the price of a house?

Student Name : Haitham Qandeel

Date : 15th Jan 2023

Github: https://github.com/haithamqandel

* Introduction

## Real estate industry is one of the most vital industry in the world. However, it can be difficult to predict the price of a house based on various factors like location, square footage, age of the building and number of rooms. In this project, we aim to develop a machine learning model that can accurately predict the price of a house by taking into account multiple factors. By using historical real estate data, we will train our model to make accurate predictions that can be used by real estate agents, buyers, and sellers alike. This can help them make more informed decisions about the price of a house in their area.

* Problem

## One of the biggest challenges in the real estate industry is to determine the right price of a property. While many factors like location, square footage, age of the building and number of rooms can have a significant impact on the price, it's not always easy to take them all into account. Additionally, the real estate market is very dynamic and the prices are changing frequently. This can lead to mispricing of the properties which can cause a lot of difficulties for both buyers and sellers. In this project, we aim to tackle this problem by developing a machine learning model that can predict the price of a house based on multiple factors.

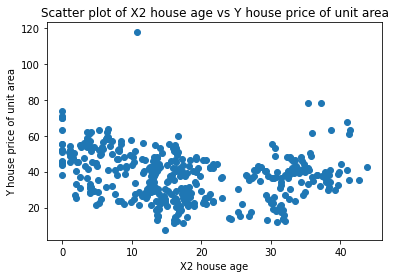
* Data set

Source : datasciencedojo

Link : <https://code.datasciencedojo.com/datasciencedojo/datasets/tree/master/Real%20Estate%20Valuation#can-you-predict-the-price-of-a-house>

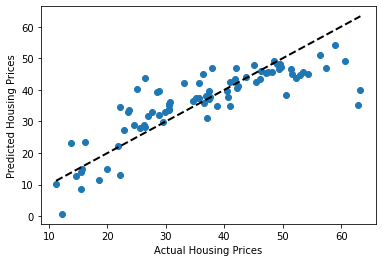
## Attribute Information

* X1 transaction date= The transaction date (for example, 2013.250=2013 March, 2013.500=2013 June, etc.)
* X2 house age= The house age (unit: year)
* X3 distance to the nearest MRT station= The distance to the nearest MRT station (unit: meter)
* X4 number of convenience stores= The number of convenience stores in the living circle on foot
* X5 latitude = The geographic coordinate, latitude (unit: degree)
* X6 longtitude = The geographic coordinate, longitude (unit: degree)
* Y house price of unit area = The house price of unit area (10000 New Taiwan Dollar/Ping, where Ping is a local unit, 1 Ping = 3.3 meter squared) for example, 29.3 = 293,000 New Taiwan Dollar/Ping
* Data visualization
  + Python . tableau ,power bi



* Model

# **Linear Regression**



* Results

the difference of the predicted values and the true values.

RMSE :7.291441874073602