

# report

September 16, 2020

## 0.0.1 Effects of Houses Characteristics and location on Their Prices

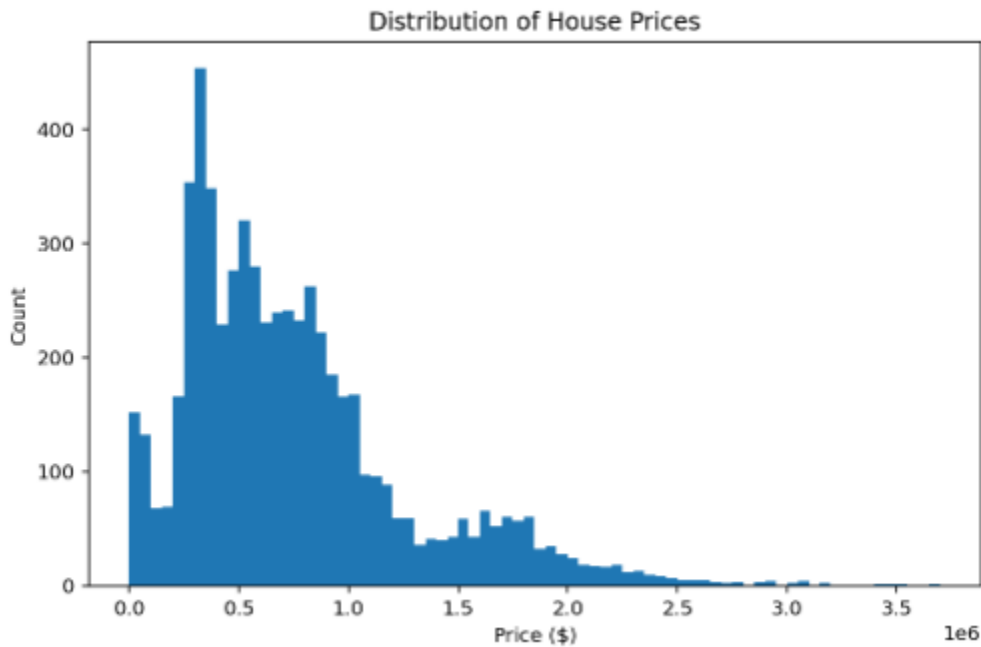
**Investigation Overview** In this investigation, I wanted to look at the characteristics of houses that could be used to predict their prices. The main focus was on the area (size) of the house, number of bedrooms, number of bathrooms and style of the house.

### Dataset Overview

The data consisted of prices and attributes of approximately 6028 houses located at three different neighborhoods (A, B and C). The attributes included the house\_id, area, number of bedrooms, number of bathrooms and style. The house\_id variable was converted from int to string type because it is solely for house identification and therefore not going to be used for computations.

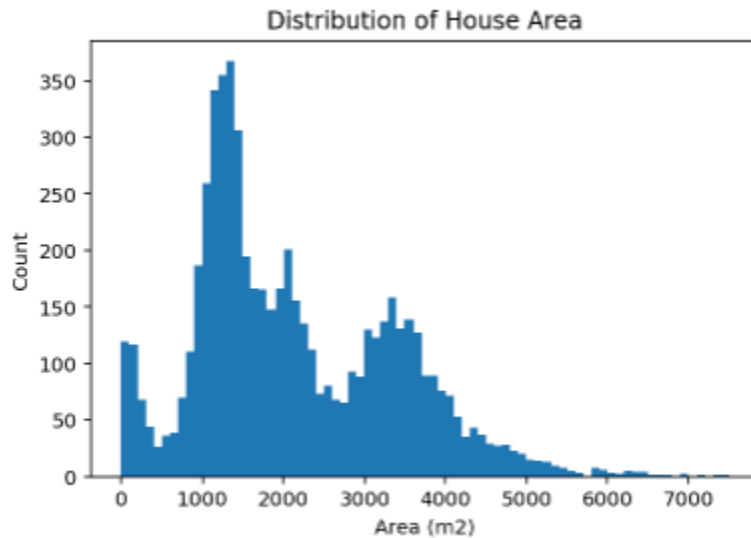
## 0.0.2 Distribution of House Prices

House prices in the dataset take on a very large range of values, from about \$12,167 at the lowest, to about \$3,684,602 at the highest. Plotted on a log-arithmetic scale, the distribution of diamond prices takes on a multimodal shape.



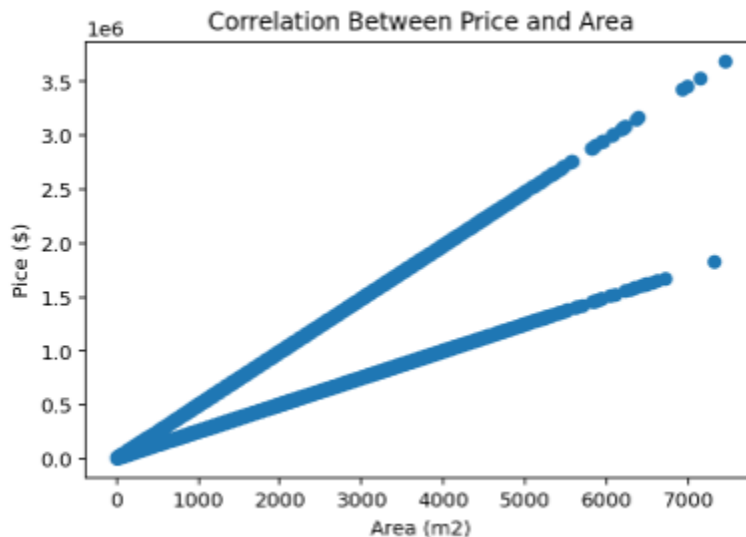
### 0.0.3 Distribution of House Areas

The distribution of the area of the houses also seems to be multimodal with most of the areas between 1000 and 2500 having the first peak between 0 and 500. The second peak between 1000 and 2500 and the last one between 3000 and 4000.



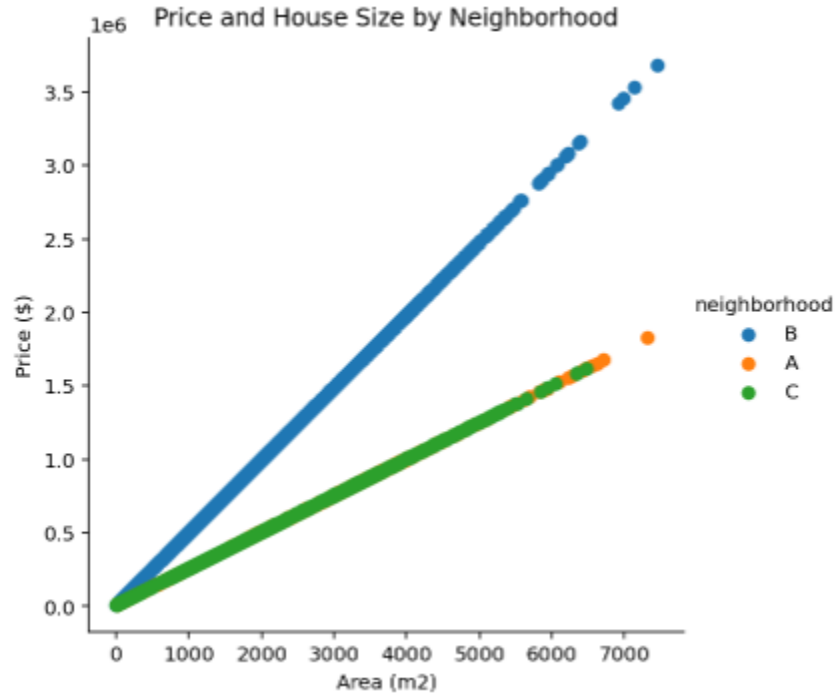
### 0.0.4 Price Vs Area (Size of the house)

Plotting the Price and Area gave almost a linear relationship. The most expensive houses has an area of 600 - 7000, although some houses of the size tend to be way cheaper than others.



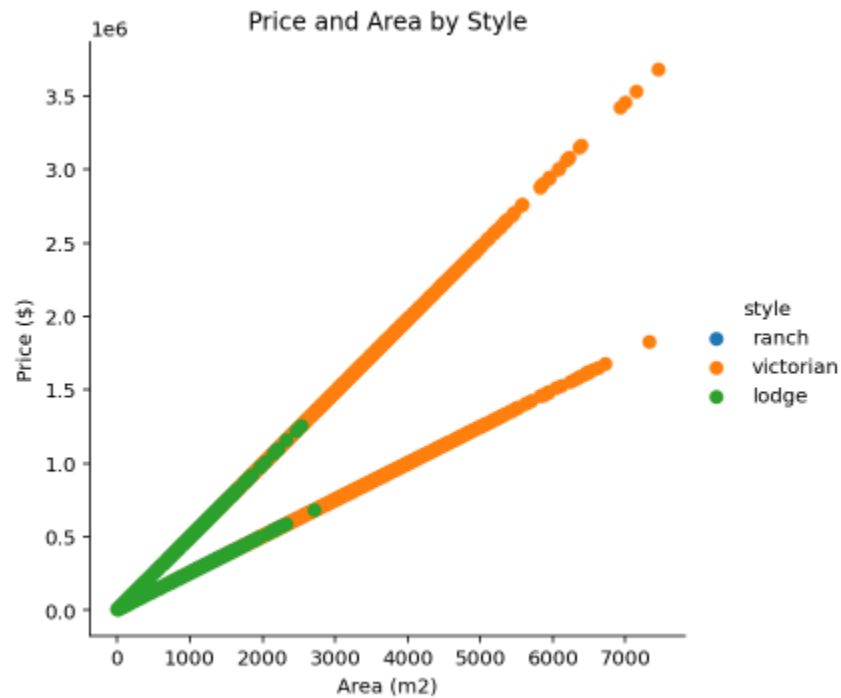
### 0.0.5 Price and House Size (Area) by Neighborhood

There exist an extreme relationship between the price of a house and the neighborhood, because as the neighborhood changes, the price of a house changes drastically. i.e the difference in price between houses in Neighborhood B is extremely higher than those in Neighborhood A and C despite having the same size (area)



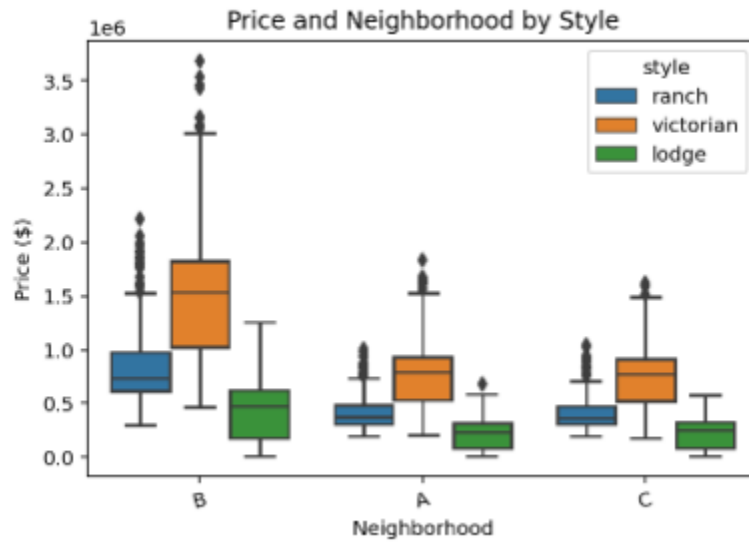
#### 0.0.6 Price and House Size (Area) by Style

The relationship between style and price is also huge, as can be seen in the figure above, all the houses with higher price from around 1.5 million to above are victorian style, and all those below 500k are not victorian.



### 0.0.7 Price and Neighborhood by Style

Both Neighborhood and Style have a strong correlation with price, the victorian style houses are much more expensive than other styles, but if compared to area of the house, victorian style in area B is much more expensive than victorian style house in neighborhood A and C as in the figure below.



[0] :