Executive Summary

The dataset is the housing prices in Auckland, New Zealand. It contains all measurements from Each house has all its factors such as location (latitude, longitude), be the number of bedrooms and bathroom, land area, 18 other variables recorded.
There are 1,051 number of observations.
There are two response variable which are the address and Auckland suburbs which the house resides in and the rest are explanatory variables. They all measure the house properties. All the measurement have summary statistics with mean, standard error.
We are interested to see if the factors influence the housing prices in Auckland.
Then, we calculate the descriptive statistics and by using visualization of the correlation between each numerical variable, we found highly co-related variables. Therefore, we tested the algorithm and the most suitable model was chosen.

Initial Data Exploration

The initial exploration of the data began with some summary and descriptive statistics. Individual Feature Statistics Summary statistics for minimum, maximum, mean, median, standard deviation, and distinct count were calculated for numeric columns, and the results taken from 1,051 observations are shown here:

Correlation and Relationships Numeric Relationships

The correlation between the numeric columns was calculated and observed in the below correlation plot.

The graph shows that the number of bathrooms, SA1 and land area of the house CV has a strong positive correlation with each other. This strong correlation can be seen on mean and standard error.

Analysis

In this analysis, an algorithm has been tested, which is simple linear regression.
The accuracy was 63.6%.

Conclusion

This analysis has shown that the house prices diagnostic can be moderately predicted from its factors of house properties. The accuracy rate is 64%.