EDA on House Sales in King County, USA

K A ADEAB -PES1201801308 NISCHAL JAIN -PES1201801298

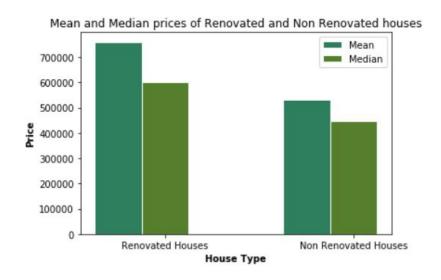
Data Cleanup

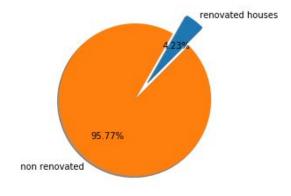
In the given dataset, we dropped rows containing 0 number of bathrooms or bedrooms.

This is because the grades implied that those houses were either under construction(i.e grade<4) or had average grades but both the number of bathrooms and bedrooms were zero.

A 0.5 bathroom implied it contained only toilet but no shower. We rounded the bathroom values in the dataset to make our plots relating to that look neater.

Comparison of Mean and Median Prices of Renovated and Non Renovated Houses





Hypothesis Testing

From the previous graphs, we infer that only a small percentage of houses were renovated and their mean/median prices were higher than the non renovated ones. We assume the given dataset as a sample and test this hypothesis.

H0: Median prices of renovated and non renovated houses are same

H1: Median prices of renovated and non renovated houses are not the same

After applying the Kruskal Wallis test in python, the statistic value was found to be 220.57. Which is greater than the critical value(3.841), hence **We reject the null hypothesis.**

Do renovated houses have higher grades?

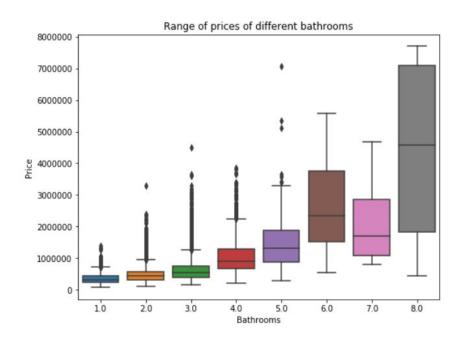
H0: There is no difference between the mean grades of renovated and non renovated ones

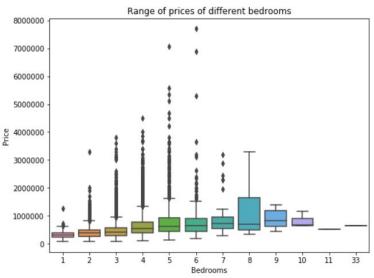
H1: There is a difference between the mean grades of renovated and non renovated ones

After applying the Mann Whitney U test in python, we found the p value as 0.0282. P value is less than our significance level(5%), hence

We reject the null hypothesis

Range of prices of different bathroom/bedrooms

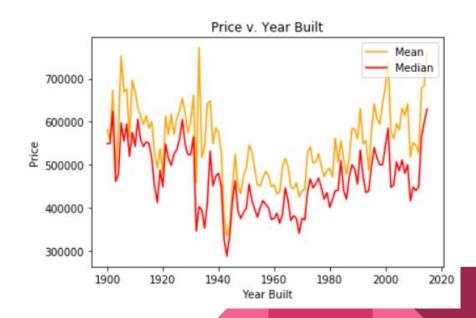




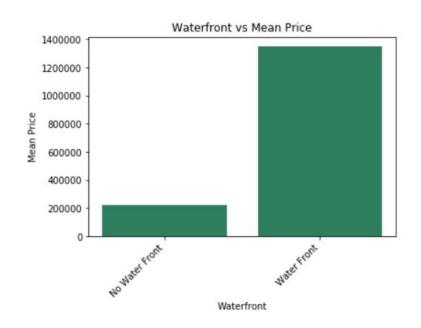
The trends in mean/median prices over the years

The median graph follows a similar pattern with mean graph.

We can observe that houses that were built from 1980's onward, show an overall positive growth in price.



Does the price of a house of a house depend on its location?

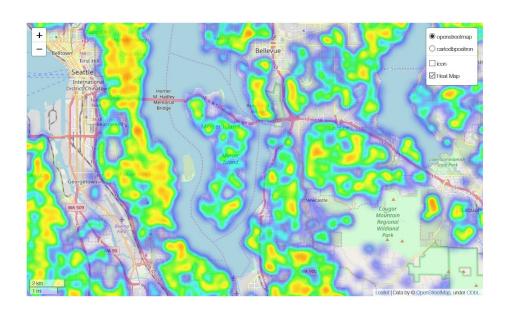


Having a waterfront is a factor of location.

Hence we'll compare the prices of houses with and without a waterfront.

Mean prices of houses with waterfront is higher.

Visualizing location vs price as a heatmap



Houses near the water bodies are either yellow or red, implying that they are more expensive.

Interior part of the land appears to be green, implying that they are relatively less expensive.