King County Housing Analysis

Summary

An analysis of the King County Housing data on Behalf of Reigman Real estate.This is to enable the company to make decisions on properties to sell or buy based off of:

- Living space of the house
- Living space of neighbouring houses
- Grade of the house

Outline

- Business Problem
- Data
- Modeling
- Regression Results
- Conclusions
- Recommendations
- Next Step

Business Problem

Reigman Real Estate is a real estate company situated in King County and they primarily deal in buying and selling of properties.In order to make well informed decisions on what type of house should be sold or bought, the company has sought out a data scientist to help them make profitable decisions based off the available data on housing sales in the region.The aim is to find which factors influence the pricing of houses and by how much

Data

Data used in this analysis came from from king county and a brief description of the columns used in the analysis is given below

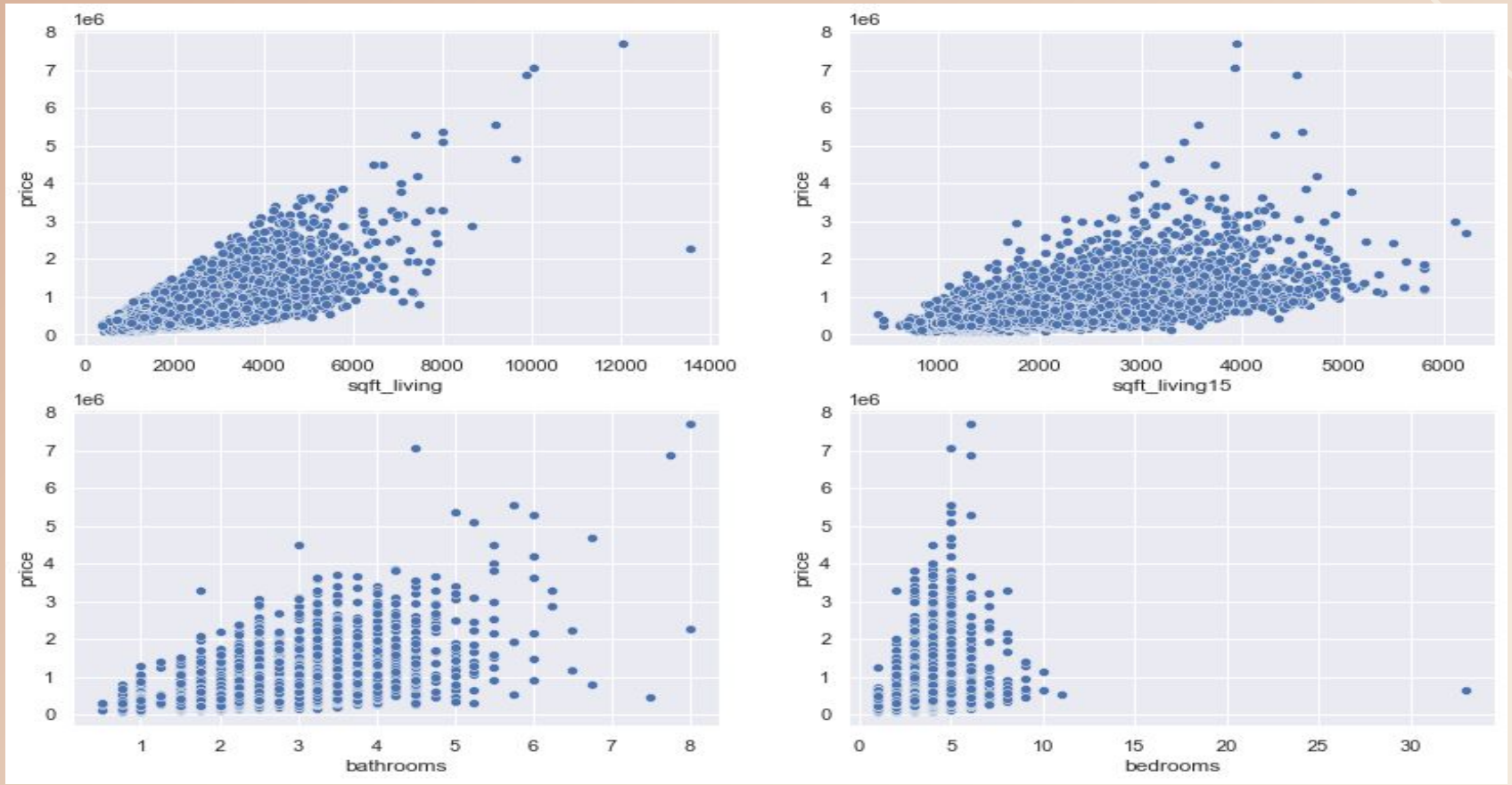
Data

- Price - Sale price (prediction target)
- Bedrooms - Number of bedrooms
- Bathrooms - Number of bathrooms
- Sqft_living - Square footage of living space in the home
- Sqft_lot - Square footage of the lot
- Condition - How good the overall condition of the house is. Related to maintenance of house.
    - See the King County Assessor Website (https://info.kingcounty.gov/assessor/esales/Glossary.aspx?type=r) for further explanation of each condition code
- Grade - Overall grade of the house. Related to the construction and design of the house.
    - See the King County Assessor Website (https://info.kingcounty.gov/assessor/esales/Glossary.aspx?type=r) for further explanation of each building grade code
- Sqft_living15 - The square footage of interior housing living space for the nearest 15 neighbors

Modeling

The purpose of modeling is to enable us to predict values based off the available data.The column we are most interested in is price and we would like to see how the other features of a house affect it.

The correlation of these features to price can be show via scatterplots


1e6
price
sqft_living
sqft_living15
bathrooms
bedrooms

Modeling

The overall model is statistically significant having an adjusted r-squared of about 60 meaning that 60% of the variance in price is covered.In predictions the model is of by $156659.

From the constant and coefficients of the column the price can be predicted by multiplying the associate value from a column with its coefficient getting the sum of all the features and their associated values and adding the constant

Regression results

- The intercept is at about $121700. This means that a house with a
- `grade` of 7(average) would sell for $121700.
- The coefficient for `sqft_living` is about $152. This means for each additional squared foot increase,
- the house costs about $152 more.
- The coefficient for `sqft_living15` is about $14. This means for each additional squared foot increase for the nearest 15 neighbours,
- the house costs about $173 more.
- The coefficients for `grade` range from about $2395000 to about -$42500
    - For a grade of "5(fair)" compared to a grade of "7(average)", we expect -$42500 price
    - For a grade of "6(low average)" compared to a grade of "7(average)", we expect -$20860 price
    - For a grade of "8(good)" compared to a grade of "7(average)", we expect +$59560 price
    - For a grade of "9(better)" compared to a grade of "7(average)", we expect +$178600 price
    - For a grade of "10(very good)" compared to a grade of "7(average)", we expect +$372100 price
    - For a grade of "11(excellent)" compared to a grade of "7(average)", we expect +$657500 price
    - For a grade of "12(luxury)" compared to a grade of "7(average)", we expect +$1192000 price
    - For a grade of "13(mansion)" compared to a grade of "7(average)", we expect +$2395000 price

Conclusion

The findings from the modeling can lead us to conclude:

- The bigger the houses living space the higher the price
- The bigger the living spaces of the neighbouring 15 houses the higher the price of the house
- The price of houses with a grade below 7(average) compared to the price of houses with a grade of 7 is lower
- The price of houses with a grade above 7(average) compared to the price of houses with a grade of 7 is higher

Recommendations

- Houses with a bigger living space should be prioritized when selling since they lad to an increase in price
- Houses in neighbourhoods with big houses should also be priority when selling or buying if they are to be sold later at a higher price
- Houses below grade 7 should not be considered for buying or selling.How ever they can be bought if there are intentions to renovate or increase living space
- Houses above grade 7 can be sold or bought

Next Step

The company should find more data on houses that include features such as:

- Renovations made on the home
- Security in the area
- Availability of health facilities, schools,shopping centres and recreational facilities
- Quality of the road
- Nearby towns

Thank You!

Email: mponzi.jm@gmail.com
GitHub: @Mponziii
LinkedIn: linkedin.com/in/James Mponzi/