King county House prices

20.4.2023

Amos kipkirui, Emilly Njue, Brian Muli, Samwel Muiru, Swaleh Athuman, Heddy Berit

Overview

- A real estate agency located in King County is looking to advice homeowners about how home renovations might increase the value of their homes, and by what amount. It is therefore crucial to look at the different house features in relation to price We will also develop a model to determine the prices of different houses with different features.

Business understanding

Introduction

- The real estate market is a dynamic and ever-changing industry, and accurate prediction of housing prices is crucial for both buyers and sellers. In order to make informed decisions, stakeholders in the real estate industry need access to reliable and comprehensive data.

The King County House Sales dataset is a valuable resource for understanding the dynamics of the real estate market in King County. This dataset contains detailed information on house sale prices including a wide range of features such as the number of bedrooms, bathrooms, square footage, location, and more.

- This dataset allows for in-depth analysis and modeling to understand the factors that influence housing prices in the region, and serves as a valuable resource for developing and testing predictive models for accurate price predictions.
- We will provide an overview of the King County House Sales dataset, including its key features, data quality, and potential use cases. We will also highlight the significance of this dataset for evaluating regression models to predict housing prices in King County, and the potential benefits it can offer to stakeholders in the real estate industry.

Business understanding cont.

To aid in making these recommendations, we will attempt to answer these questions:

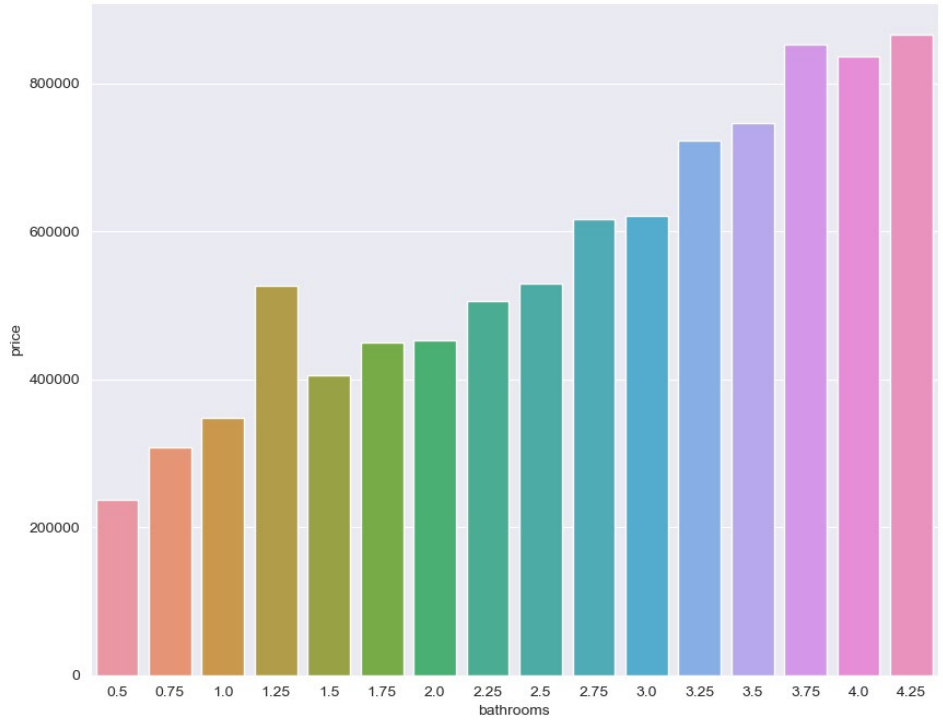
1. Determine what are the key factors that significantly impact housing prices in King
2. How does the number of bedrooms, bathrooms, grade and square footage of a house correlate with its sale price in King County?
3. Predict House sale prices given houses specifications

Data exploration

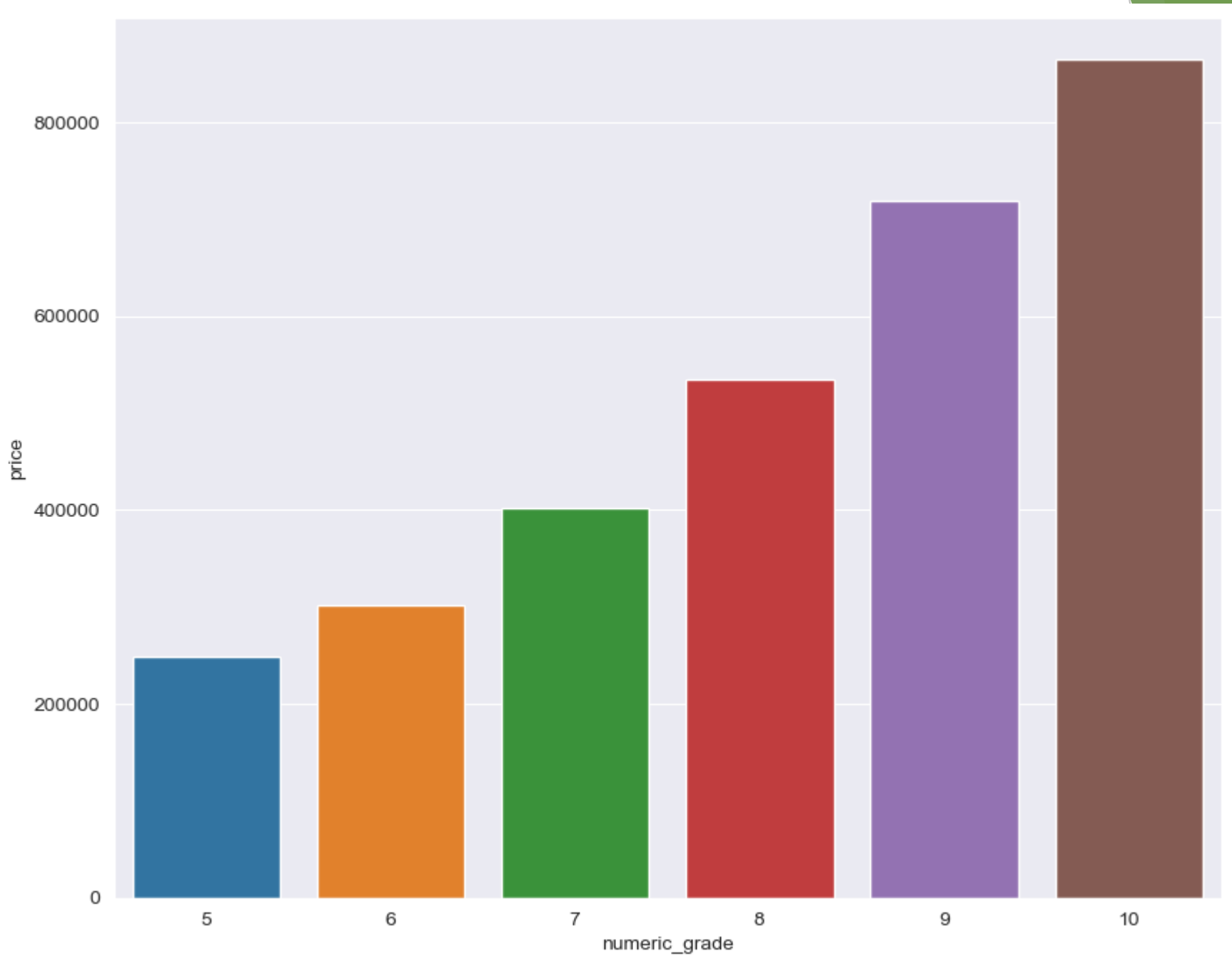
- We used king county house sales datasets; kc_house_data.csv

Finding the relationship between bathrooms and price

Also, there's a relative increase in price to number of bathrooms.



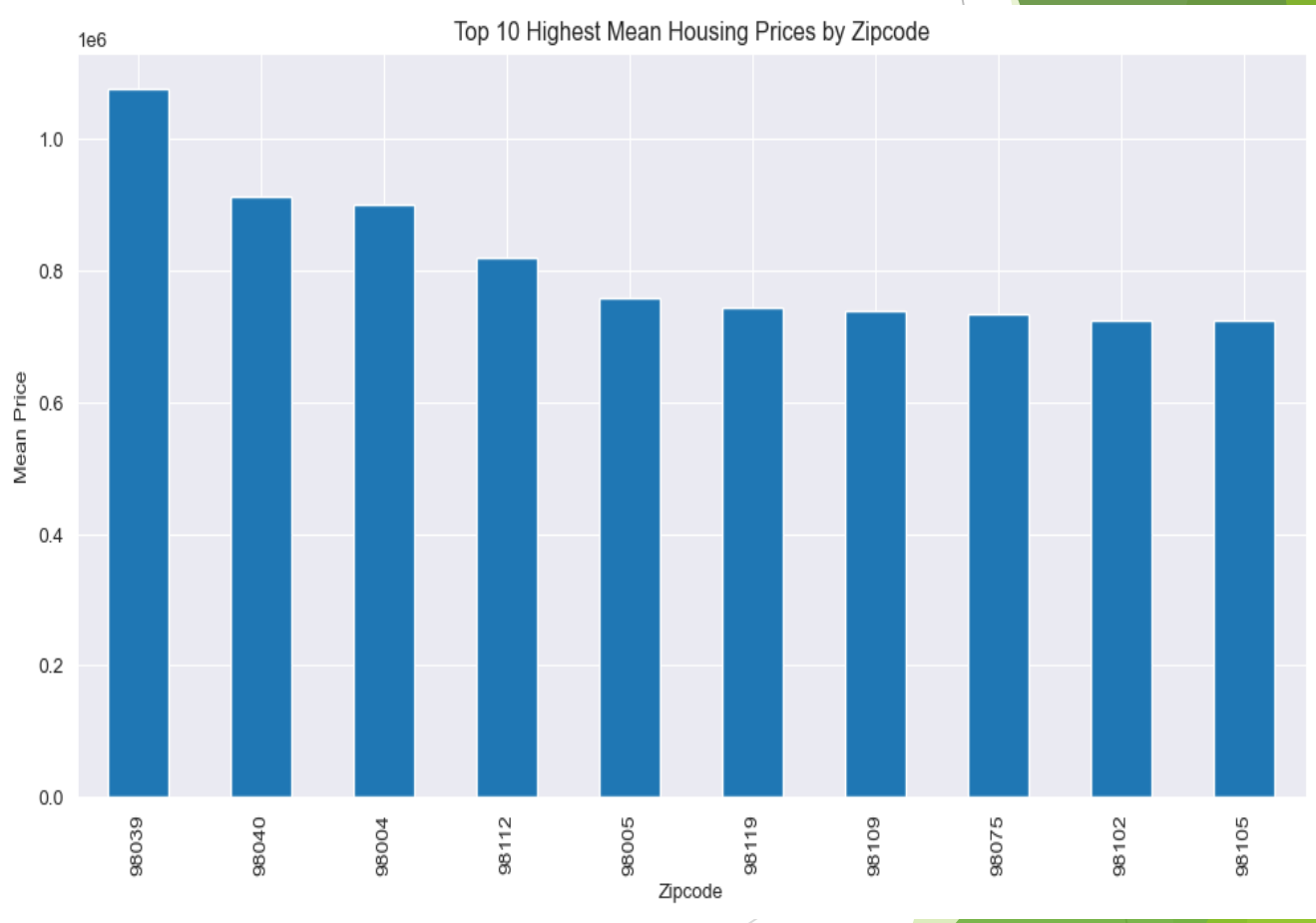
Finding the relationship between house grades and the corresponding prices


800000
600000
400000
200000
0
price
5
6
7
8
9
10
numeric_grade

From the above graph we can see that an higher house grades had higher prices and so the homeowner should ensure that his/her house is of high quality as this will attract high grading hence high prices.

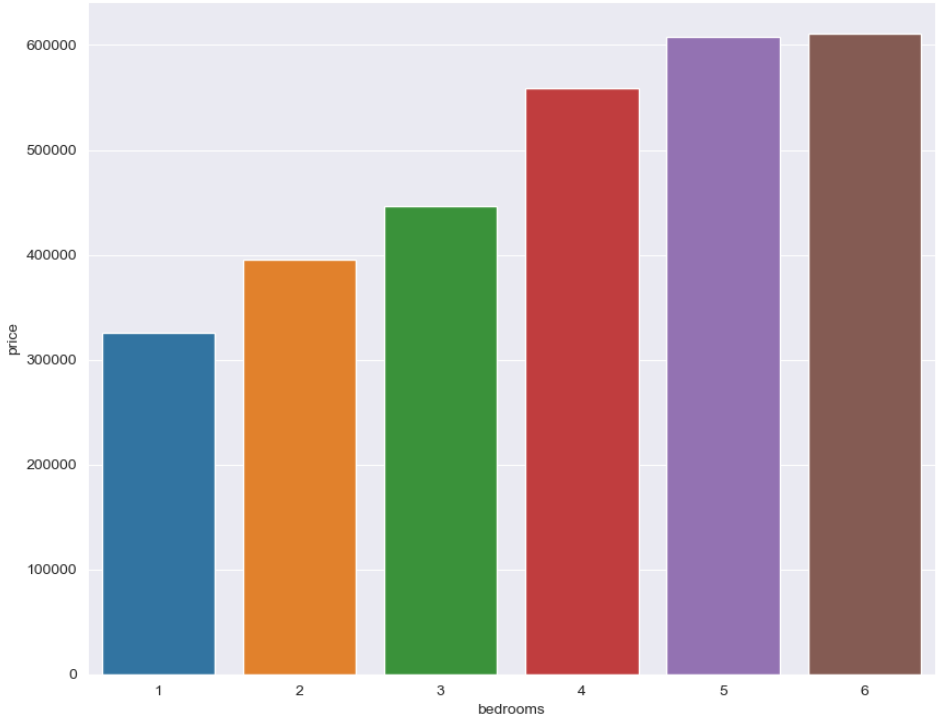
Finding the relationship between zip code(locations) and price

From the graph, we can see that there are certain locations represented by zip codes that had higher prices as compared to others. The buyer should therefore buy a house in these locations.



Finding the relationship between bedrooms and price

An increase in number of bedrooms in a house results to a corresponding increase in the price of a house. The homeowners(sellers) should consider increasing the number of bedrooms during the renovation process in order to demand a higher price of the house.



Insights;

We threw the following insights from our modeling and the exploratory data analysis we performed on our data:

1) This project aimed to develop a model to predict housing prices in King County based on various features such as square footage, number of bedrooms and bathrooms, grading and location(zipcode).
2) Firstly, we performed exploratory data analysis (EDA) and found that the price of houses was positively correlated with the square footage and the grade of the house. We also discovered that the location of the house had a significant impact on the price.
3) We then developed multiple linear regression models to predict the price of the house, and we
found that the model that included square footage, grade, bedrooms, bathrooms and zipcode as
the predictors performed reasonably well, explaining about 48% of the variance in price.

Insights cont.

3. We then developed a classification model to predict whether a house is expensive or affordable based on features such as square footage, number of bedrooms and bathrooms, grade and location. We used logistic regression and found that the model performed well, with an accuracy of 73%.

4. Based on the results, we can conclude that square footage and the numeric grade of the house are significant predictors of housing prices in King County. Additionally, we can use the classification model to predict whether a house is expensive or affordable based on the features of the house.

Recommendations

From the above insights, we recommend the following:

1) Homeowners should focus on highlighting the size of their homes if they want to attract buyers in higher price categories. This means showcasing the square footage of the property.
2) Buyers who are seeking higher-end properties should pay closer attention to the location and zip code of the properties they are considering. Properties located in certain zip codes are more
likely to be expensive, so buyers should research the local real estate market and focus on properties in desirable areas.

Recommendations contd.

3. Houses with a higher "numeric_grade" are more likely to fall into the higher price category.
Therefore, it would be recommended for homeowners to consider investing in improvements to
their home's grade in order to potentially increase its value.

4. Buyers who are looking for higher-end properties may not necessarily prioritize the number of
bathrooms and bedrooms. Therefore, sellers should be aware that adding additional bathrooms
or bedrooms may not necessarily increase the value of the property.

Next steps

- Further research and analysis can be done to improve the accuracy of the models and to gain a
better understanding of the factors that influence housing prices in King County.