King County House Prices

Analysis - Prediction - Recommendations

Created by: Manuel Klein
Date: 11/04/2019

King County House Prices

Task:

- Analyzing the influencing factors on house prices in King County, Washington, USA
- Creation of a model for predicting house price
- Deriving recommendations for house sellers and buyers

Created by: Manuel Klein
Date: 11/04/2019

King County House Prices

Input Data:

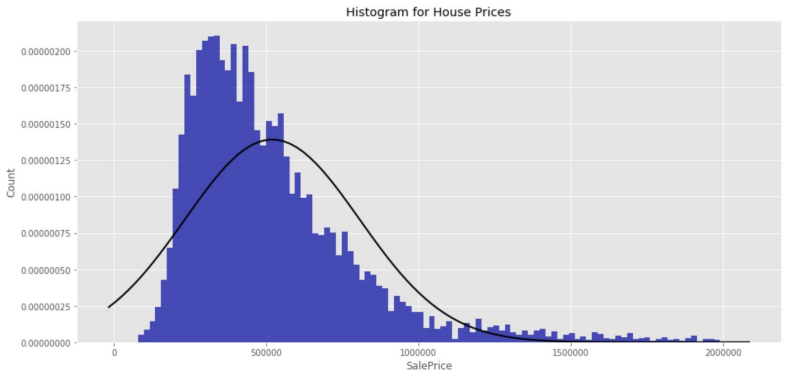
- .csv file from the statistics department
  - 20 features
  - 21597 samples
- Feature explanations file

Created by: Manuel Klein
Date: 11/04/2019

Features contained in the raw data:
* **id** - unique identified for a house
* **dateDate** - house was sold
* **pricePrice** - is prediction target
* **bedroomsNumber** - number of Bedrooms/House
* **bathroomsNumber** - number of bathrooms/bedrooms
* **sqft_livingsquare** - footage of the home
* **sqft_lotsquare** - footage of the lot
* **floorsTotal** - floors in house
* **waterfront** - House which has a view to a waterfront
* **view** - Has been viewed
* **condition** - How good the condition is (Overall) --> Details below
* **grade** - overall grade given to the housing unit, based on King County grading system --> Details below
* **sqft_above** - square footage of house apart from basement
* **sqft_basement** - square footage of the basement
* **yr_built** - Built Year
* **yr_renovated** - Year when house was renovated
* **zipcode** - zip
* **lat** - Latitude coordinate
* **long** - Longitude coordinate
* **sqft_living15** - The square footage of interior housing living space for the nearest 15 neighbors
* **sqft_lot15** - The square footage of the land lots of the nearest 15 neighbors

King County House Prices

Overview:



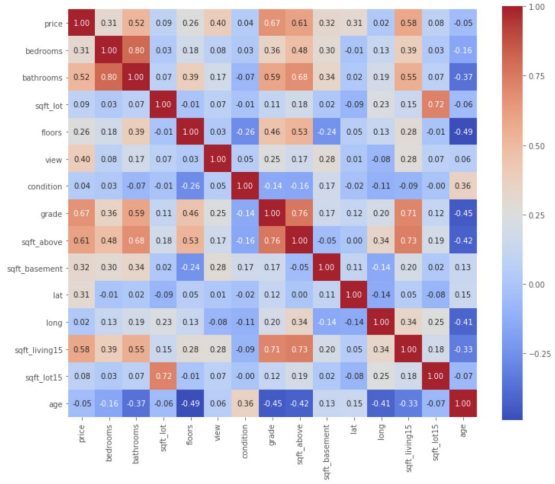
Mean: 541,449 USD Std: 370,901 USD

Created by: Manuel Klein
Date: 11/04/2019

For better readability, the Histogram has been reduced by approximately the highest 1% of house prices. Thus outliers are not shown in the histogram.
Min = 78,000 USD Max = 7,700,000 USD

King County House Prices

Feature correlation:



Created by: Manuel Klein
Date: 11/04/2019

The feature correlation matrix shows how the features are correlated to each other. Before creation, the data has been cleaned. Moreover, categorial variables have been removed for this matrix. These variables have been analyzed seperately.

King County House Prices

10 most important features:

1. Square footage of house apart from basement
2. Grade based on King County grading system (categorial)
3. Zip code of the house location (categorial)
4. The square footage of interior housing living space for the nearest 15 neighbors
5. Number of bathrooms
6. Number of prospective buyers that have viewed the house
7. Square footage of the basement
8. Latitude coordinate of the house position
9. Number of bedrooms
10. Number floors

Created by: Manuel Klein
Date: 11/04/2019

The 10 most important features for the regression contain features from the correlation matrix as well as categorial features.

King County House Prices

Multiple Linear Regression:

OLS Regression Results

Dep. Variable:	price	**R-squared:**	0.811
Model:	OLS	**Adj. R-squared:**	0.811
Method:	Least Squares	**F-statistic:**	943.2
Date:	Sun, 03 Nov 2019	**Prob (F-statistic):**	0.00
Time:	12:05:48	**Log-Likelihood:**	-2.5696e+05
No. Observations:	19164	**AIC:**	5.141e+05
Df Residuals:	19076	**BIC:**	5.148e+05
Df Model:	87		
Covariance Type:	nonrobust		

Created by: Manuel Klein
Date: 11/04/2019

The created multi-linear regression model is able to explain 81% of the variance of the house prices.

King County House Prices

Recommendations for house sellers:

- Increasing the grade of the house
- Transforming living space into bathroom

Created by: Manuel Klein
Date: 11/04/2019

1. The King County grading system is highly correlated to the house prices. Our recommendation: Before selling your house, put some work into it so that it is higher graded afterwards. If you cannot do it by yourself, find some craftsmen to get this job done for you. This significantly increases the selling price of the house.
2. If the number of bathrooms in your house is lower than the average bathrooms per bedrooms rate, you should transform a small amount of the living space into another bathroom. The number of bathrooms clearly correlates to the selling price of the houses.

King County House Prices

Recommendations for house buyers:

- Investing into a house with a lower grade
- Keeping some distance to the water
- Avoid zip codes where prices are disproportionally high

Created by: Manuel Klein
Date: 11/04/2019

1. The King County grading system is highly correlated to the house prices. Our recommendation: Buy a house that is graded a bit below your personal demands. Put some work into it so that the house is higher graded afterwards. If you cannot do it by yourself, find some craftsmen to get this job done for you.
2. The closer a house is located at the water, the higher the price. If living close to the water is not important for you, avoid it!
3. The house price is highly dependent on the zip codes. In order to find a cheap house, certain zip codes shall be avoided.

King County House Prices

Next steps:

- Testing the prediction quality of the model
- Analyzing outlier impact
- Cluster analysis of zip codes
- Deeper feature engineering

Created by: Manuel Klein
Date: 11/04/2019

Test data should be used to test the prediction quality of the model. Should be done in order to increase the precision of the model.
Outliers have not been removed from the data so far. Although the impact is not expected to be high, removing them might also increase the quality of the model a bit.
The zip-codes might be helpful to group and cluster the given dataset, which is likely very helpful with regard to make more precise house price predictions.
For now, all features were just analyzed completely each. In order to get deeper insights into dependencies of the house prices on the given features, it might be helpful to split the given data for each feature (e.g. less than 3 bathrooms, 3-5 bathrooms, more than 5 bathrooms).

King County House Prices

Questions?

Created by: Manuel Klein
Date: 11/04/2019

King County House Prices

Thank you!

Created by: Manuel Klein
Date: 11/04/2019