King County house sales Analysis

Ly Nguyenova, Stuti Singh

TABLE OF CONTENTS

01

ABOUT US

Here you could describe the topic of the section

02

Data ANALYSIS

Here you could describe the topic of the section

03

GOALS & STRATEGY

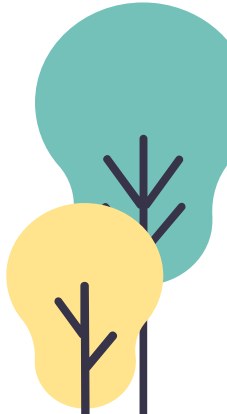Here you could describe the topic of the section

04

Modeling

Here you could describe the topic of the section

05

Findings

Here you could describe the topiof the section

About us

Two freelance data scientists that were hired by a real estate agency who wants to open a new office in King County.

The agency has given us a task of analysing the housing market in the County. Therefore, our task is to identify the factors affecting price in this area and to create a model that can help them predict the price.

King County

R EAL ESTATE

Data Analysis

- Handling of null values, extraneous values and outliers available in datasets
- Tried to find out variables affecting house price - waterfront, bedrooms, bathrooms, sqft_living, sqft_above, sqft_living15, grade
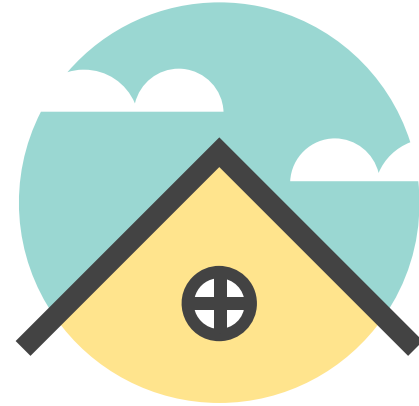- Simple linear regression and multiple linear regression model to predict price

King County

REAL ESTATE

Goal and Strategy

- Correlation between price and other factors
- Location as a factor
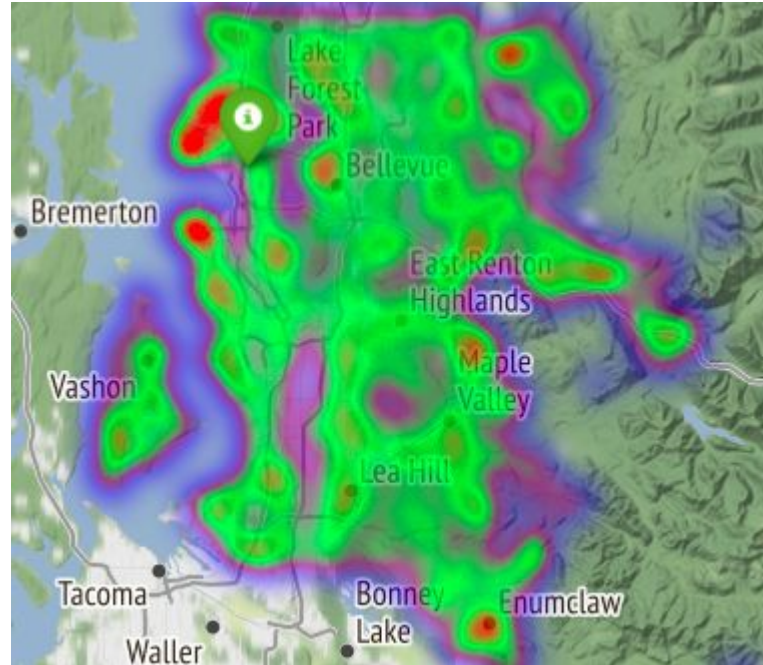- The most prominent factors
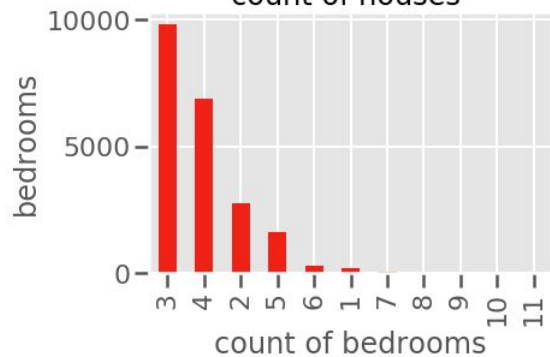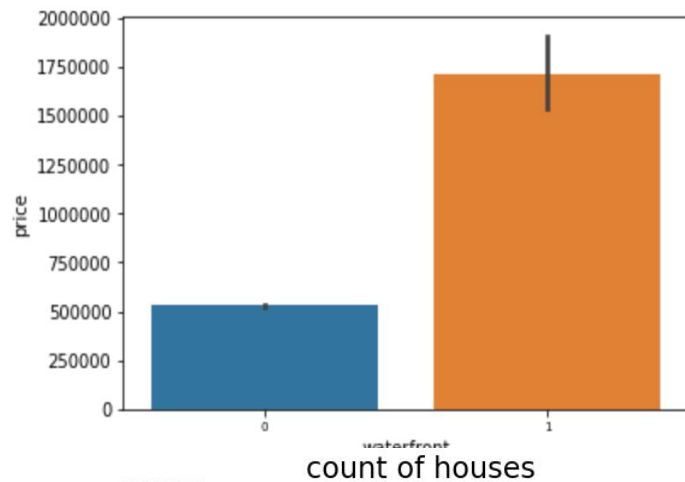
King County
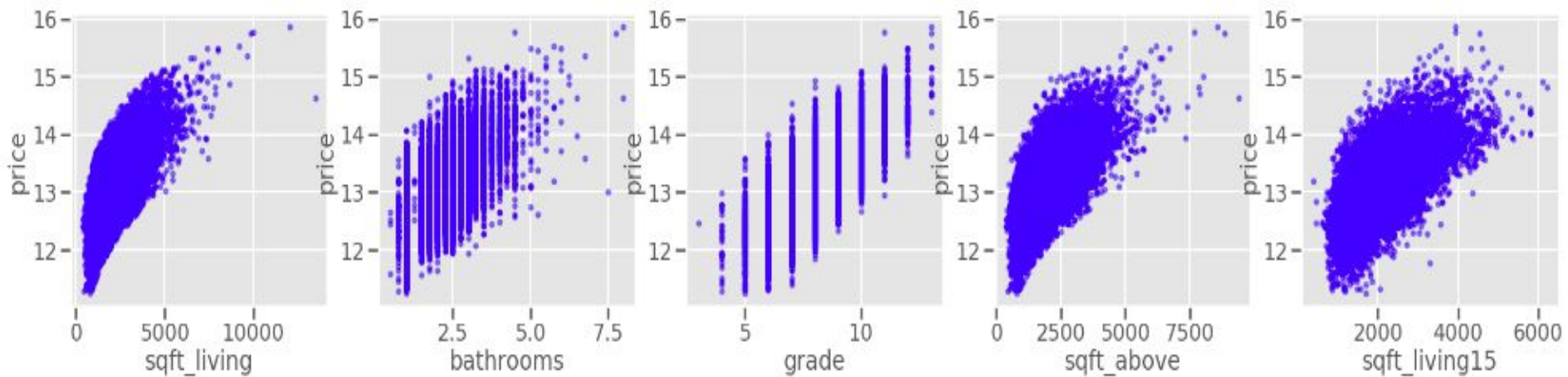
REAL ESTATE

Red spots - high price

The most expensive areas - near Seattle and the coast

```
0    21451
1      146
Name: waterfront, dtype: int64
```


	sqft_living	sqft_above	sqft_living15
sqft_living	1.000000	0.876448	0.756402
sqft_above	0.876448	1.000000	0.731767
sqft_living15	0.756402	0.731767	1.000000

Factors affecting price

Multiple Linear Regression - OLS

OLS Regression Results

Dep. Variable:	price	R-squared:	0.556
Model:	OLS	Adj. R-squared:	0.556
Method:	Least Squares	F-statistic:	8997.
Date:	Wed, 04 Dec 2019	Prob (F-statistic):	0.00
Time:	14:08:24	Log-Likelihood:	-8034.7
No. Observations:	21596	AIC:	1.608e+04
Df Residuals:	21592	BIC:	1.611e+04
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	11.1707	0.019	597.691	0.000	11.134	11.207
grade	0.1888	0.003	58.607	0.000	0.182	0.195
bathrooms	-0.0187	0.005	-3.858	0.000	-0.028	-0.009
sqft_living	0.0002	4.69e-06	48.306	0.000	0.000	0.000
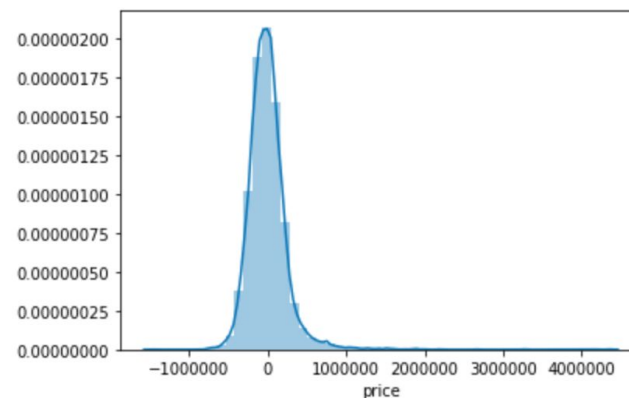
Recommendations

Houses with waterfront are more expensive

The higher grade , the higher price

The closer to the center and the northern coast, the higher price

The bigger living area square footage, the higher price

Houses with three bedrooms are in demand - higher price for houses with three bedrooms