

A decorative graphic on the left side of the slide consisting of two overlapping parallelograms. The front one is blue and the back one is a light green color. They are positioned diagonally, with the blue one in front of the green one.

# Predicting Home Prices: King County, WA

Andre Yu Tiamco , Cameron DeArman

# King County , WA

- Population : 1,931,249
- Area : 2,131 miles<sup>2</sup>
- 14th most populated county in the U.S.
- Notable Cities:
  - Seattle
  - Bellevue



# King County, WA



- King County was Originally named after William R. King
- William R. King had just been elected as Vice-President under Franklin Pierce
- In February of 1986, King County Council approved a motion to rename the county after Martin Luther King Jr.



# King County Housing Data

- In the years of 2014 and 2015 we took 21,584 homes that were sold with other collected data.
- From this data we chose to look at price of the houses when they were sold and use other data to predict future prices in King County.
- Our Most valuable features:
  - Zip Code
  - Square Foot Living
  - Grade



# Zip Code

- When then discovered a website which contained all of King County's Cities by zip code.
- Each zip code had a link associated with it which led us to finding median income, median home price ,comfort index, median age ,and the unemployment rate.
- Through some data scraping we were able to use these data points, specifically median price of homes, in our model .
- By dividing our data set into their associated zip codes we were able to boost model's accuracy my 10%

# Square Foot Living

- Square foot living, as defined by sfgate, is the total area of the house on all floors and subtract the area of that which is not heated.
- This narrows down most homes to that of main floors and floors that are typically used daily.
- Before adding other statistics, square foot living was our most correlated data point with price.



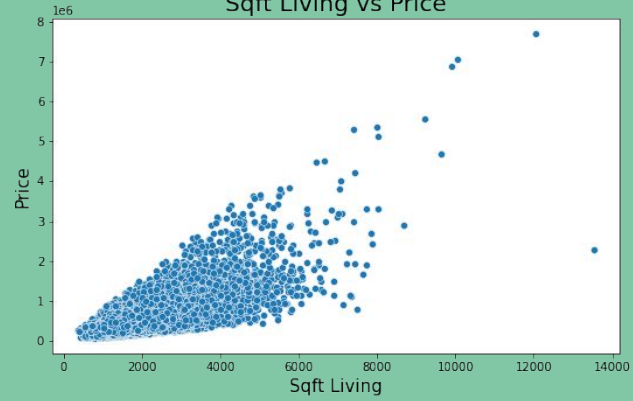


# Grade

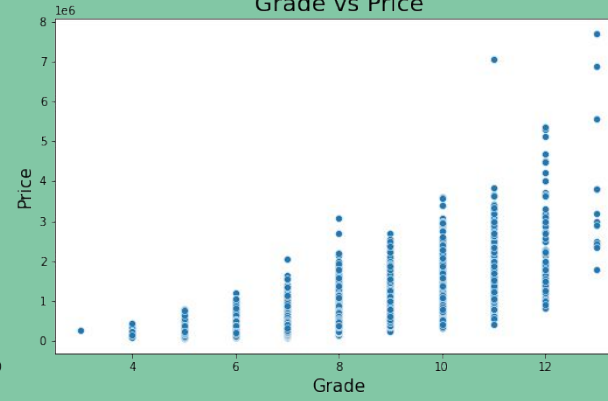
- For the definition of grade we looked to the King County website.
- Grade in King county represents the construction quality of improvements.
- The grades run on a 1 - 13 scale.
- In our model we dealt with houses on a 3 - 13 scale.

A+

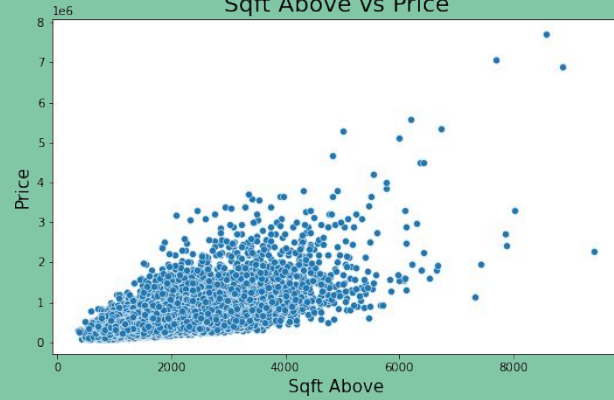
Sqft Living vs Price



Grade vs Price



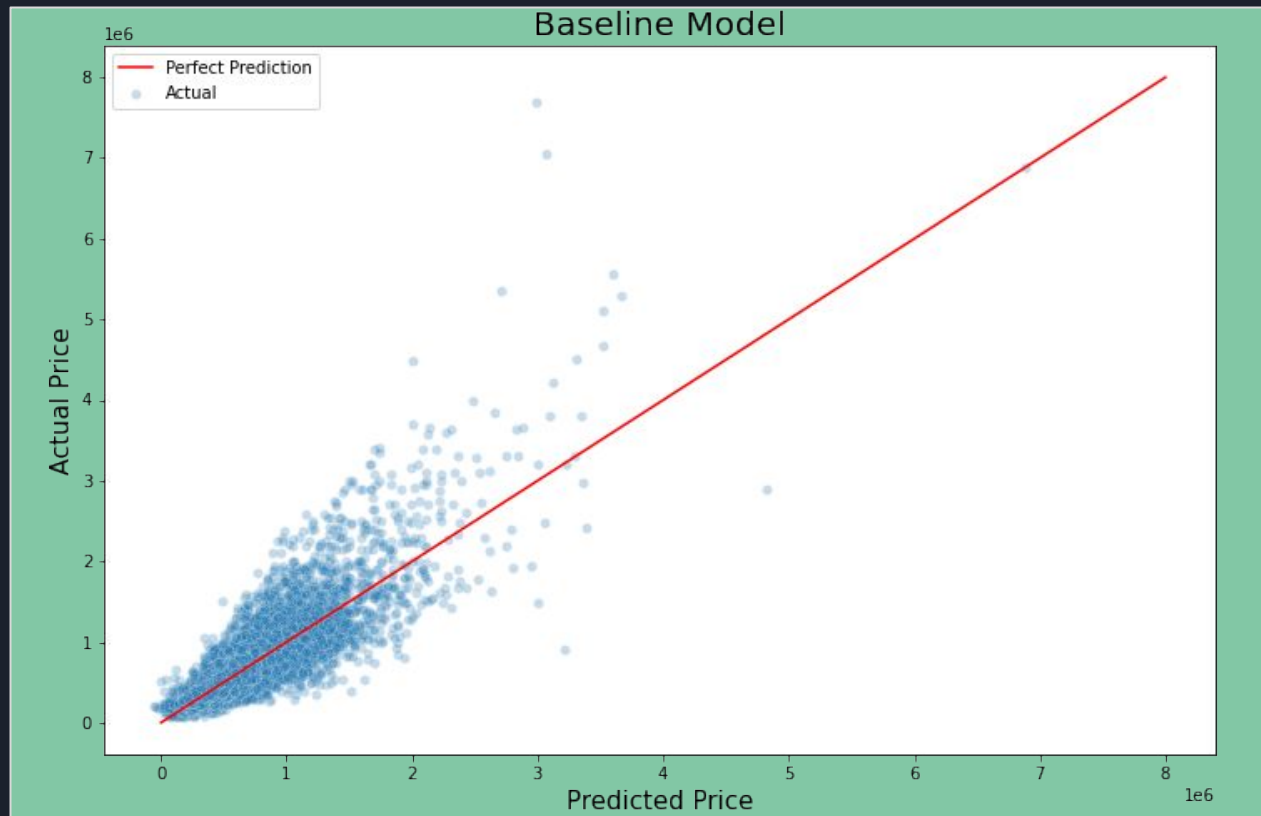
Sqft Above vs Price



Sqft Living15 vs Price

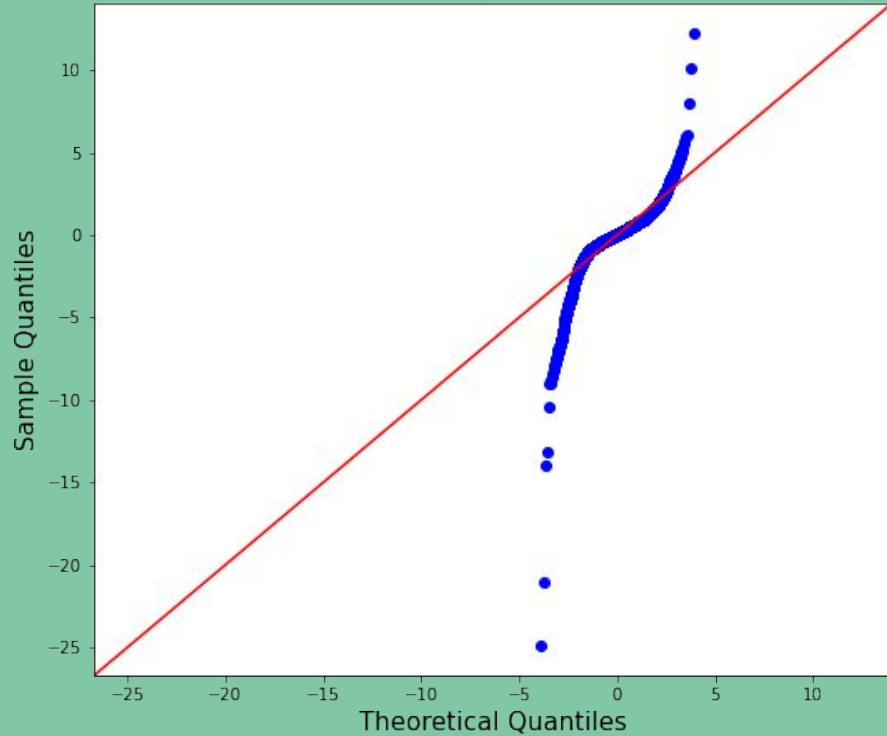




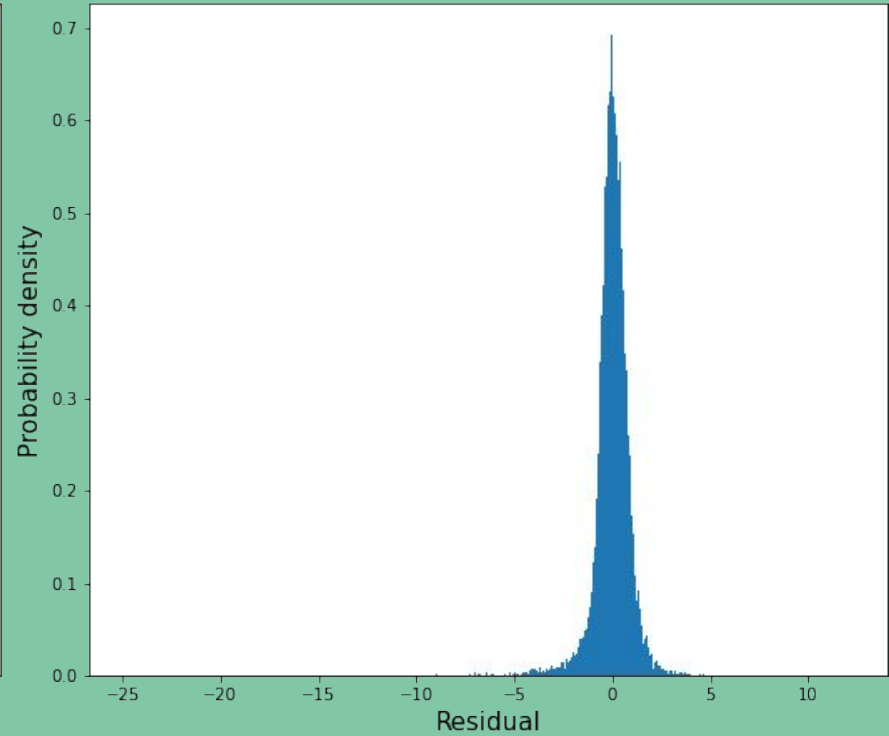


$$R^2 = 0.75$$

Baseline Standardized Residuals  
Q-Q Plot



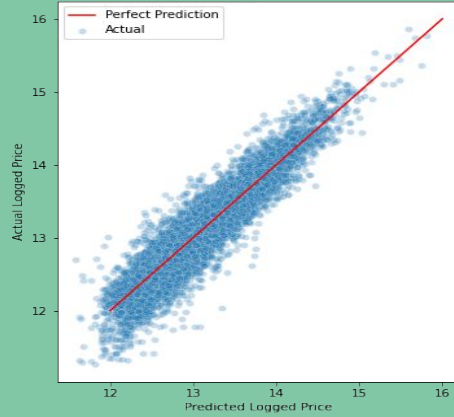
Baseline Standardized Residuals  
PDF



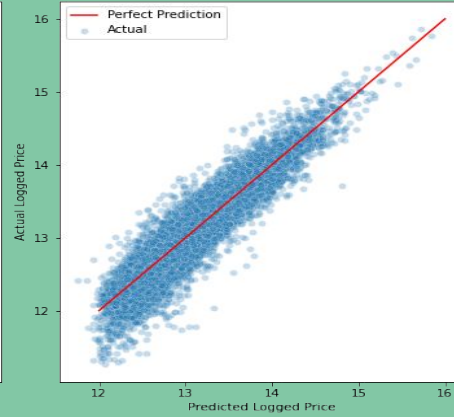
Q-Q does not follow normal line well.

PDF looks normal but has a very high peak.

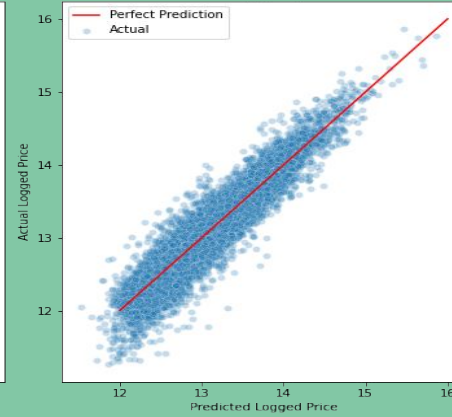
Minimize RMSE  
of Logged Price



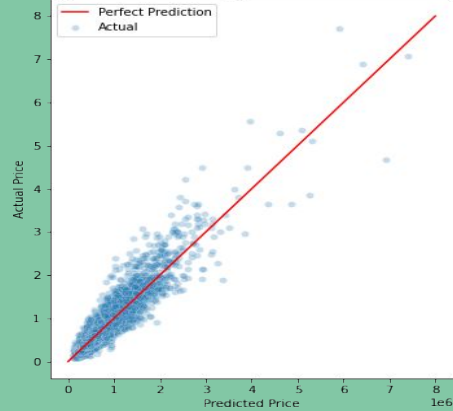
Minimize RMSE  
of Price



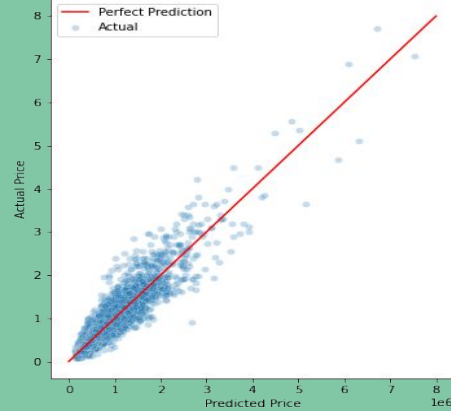
Maximize R-squared  
with logged price



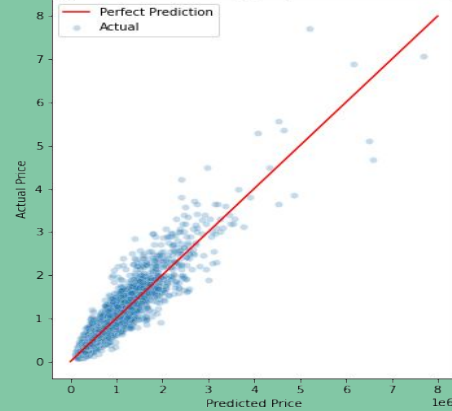
Minimize RMSE  
of Logged Price

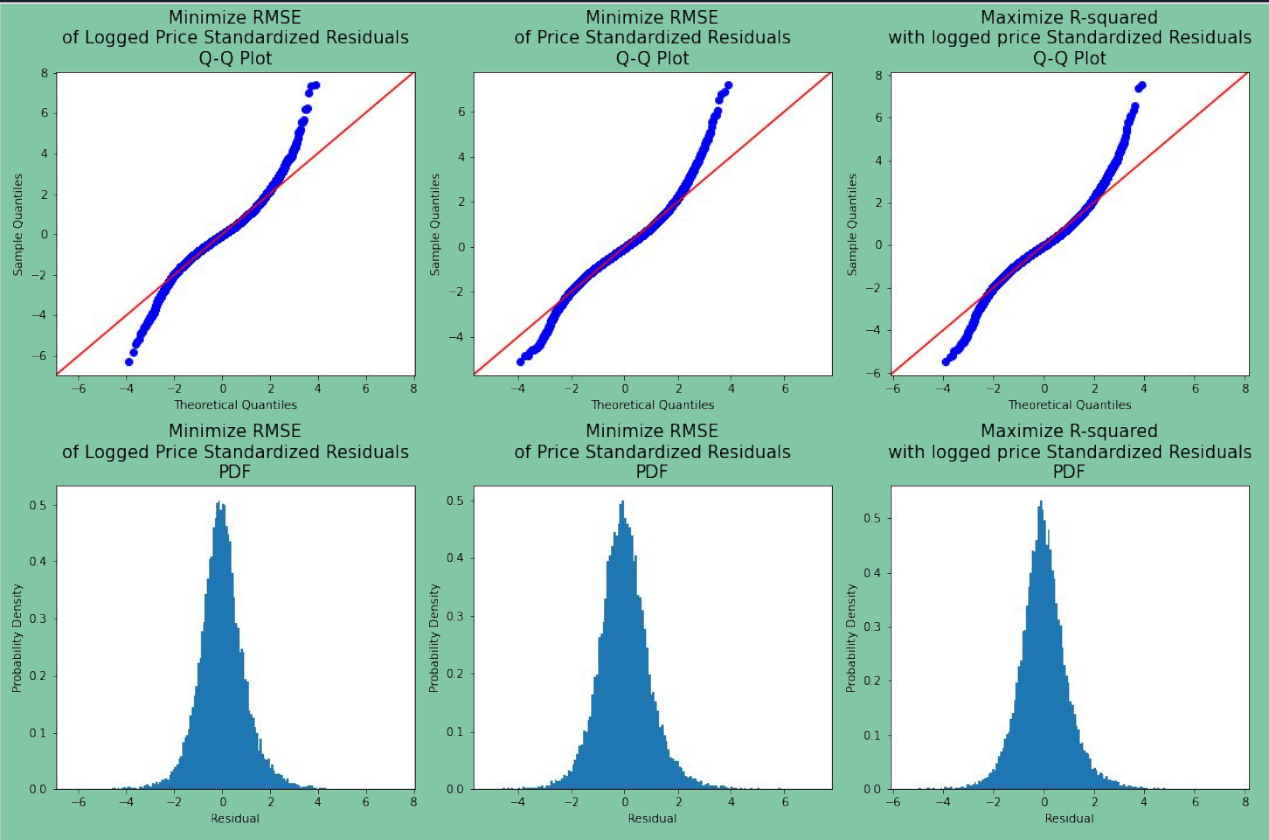


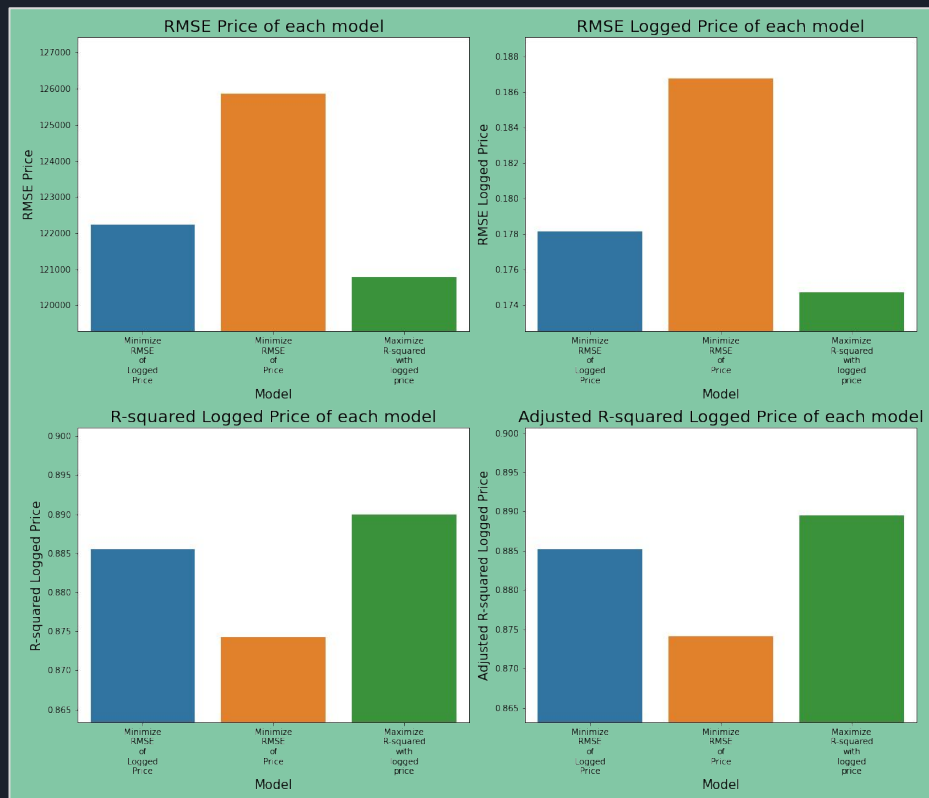
Minimize RMSE  
of Price



Maximize R-squared  
with logged price







	Features	RMSE Price	RMSE Logged Price	R-squared Logged Price	Adjusted R-squared Logged Price
Minimize RMSE of Logged Price	51	121903.414062	0.178138	0.885304	0.885032
Minimize RMSE of Price	36	125381.377881	0.186576	0.874285	0.874075
Maximize R-squared with logged price	72	121019.805354	0.174665	0.889993	0.889625

QUESTIONS

