

# King County Real Estate Analysis

## Summary

Analysis of King County Real Estate data to predict house prices.

# Outline

- Business problem
- Data & Methods
- Results
- Conclusions

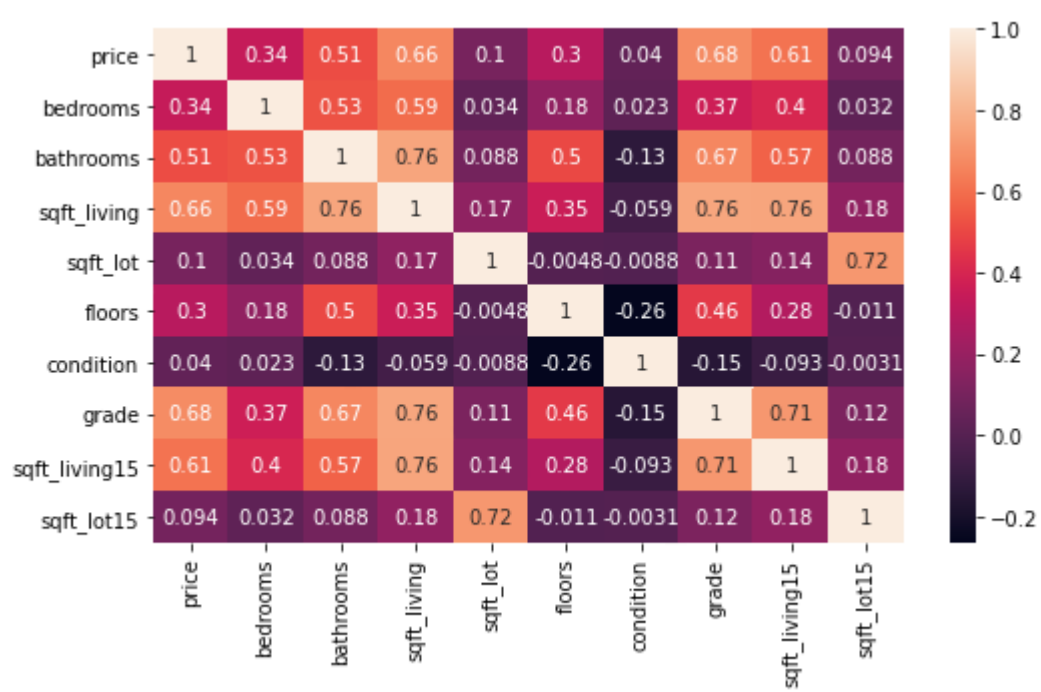
# Business Problem

- This information could be used by real estate agencies to evaluate what a property could be currently sold at within King County.
- This is would be ideal to get the best deal for both the agency and their client.

# Data & Methods

- The data used in this analysis is from real estate sales of King County.
- This dataset contains relevant information on homes sold in 2014-2015 and their details so there are a good set of variables to base the models on.
- I built linear regression models for raw data, scaled data and log transformed data to compare them and try to ascertain the most accurate predictive model.

# Results



- Price has a strong correlation with sqft\_living.
- Grade also has a strong correlation, which is the government's assessment on house quality (higher is better). So, the correlation is explained there.

# Results

	Raw data	Scaled	Log transform
Adjusted R-squared:	0.5646181	0.5234166	0.5633075

- This is the  $r^2$  for all 3 iterations and you can see that scaling the data made very little or was even a little detrimental to it.

# Conclusions

- Real estate agencies can measure house features to estimate the selling price and advise sellers.
- Houses that offer more space capacity and living quality have higher prices.

## Moving Forward

- These models need more work.
- Adding more independent variables like location will render a more accurate model.