King's County
House Sales Price Model

Problem Statement

- Data driven insights
- Model to predict house prices

In this project, we looked at the the various features that influence the price of a home in King's County. We analyzed these features, developed and tested models that would predict the price of the home

Business Value

- Better understanding of house buyer preferences.
- Assisting in house valuations by providing price predictions.

This project helps understand buyer preferences and improves the real estate market knowledge in King's County. Finally by having a model capable of predicting house prices, we can improve on our suggested sales price.

Methodology

- Obtain and analyze data.
- Investigate house features and gain insights.
- Build prediction model.

The data used for this project consists of around 21,000 prices for house sales in King's County. The first was to clean this data to ensure it is fit for modelling. We then looked at which features would provide impactful insights and built a model capable of predicting house prices.
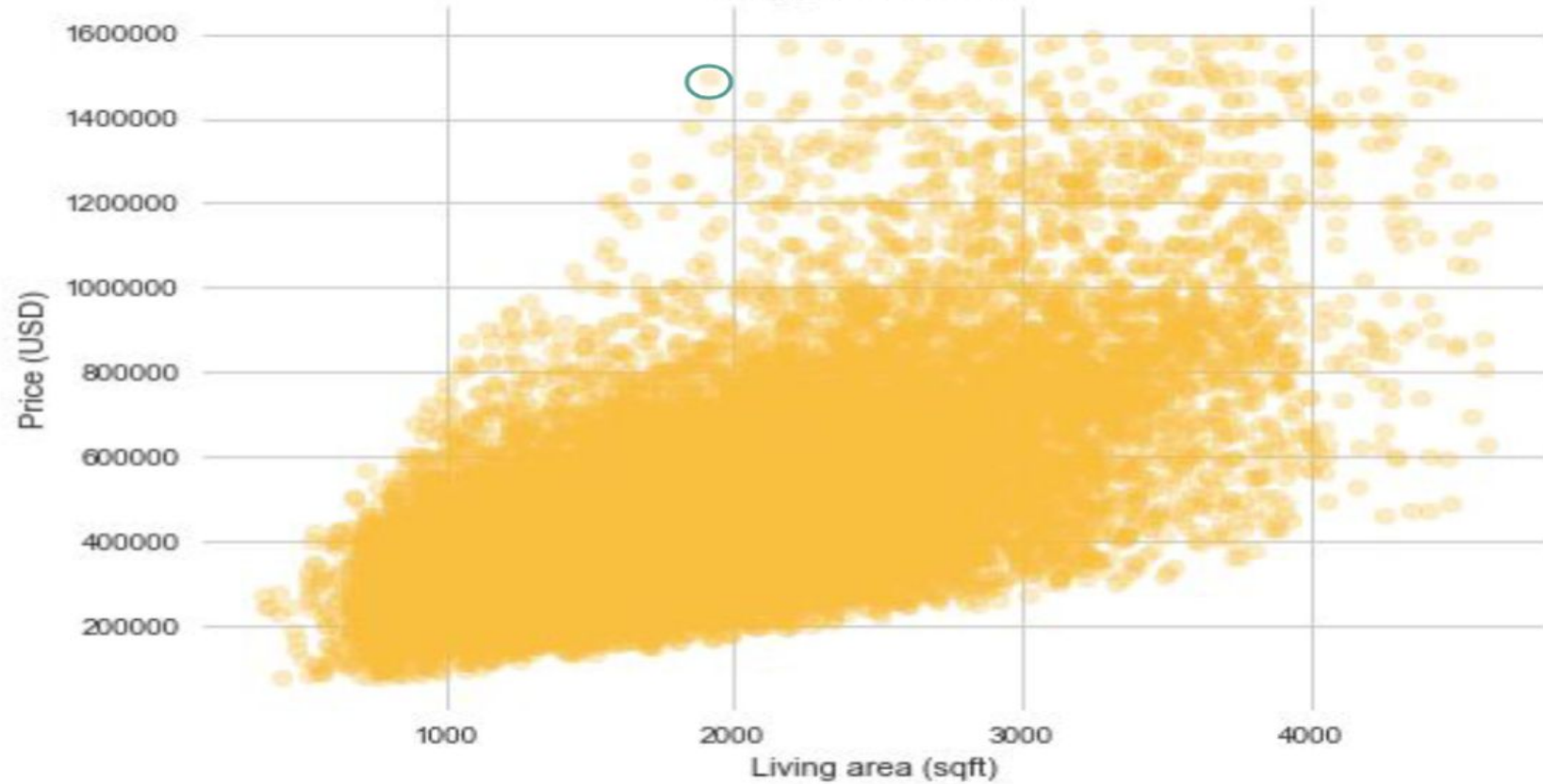
House Features

I looked at house features to see which factors would drive the price up and also which are the best predictor of price.
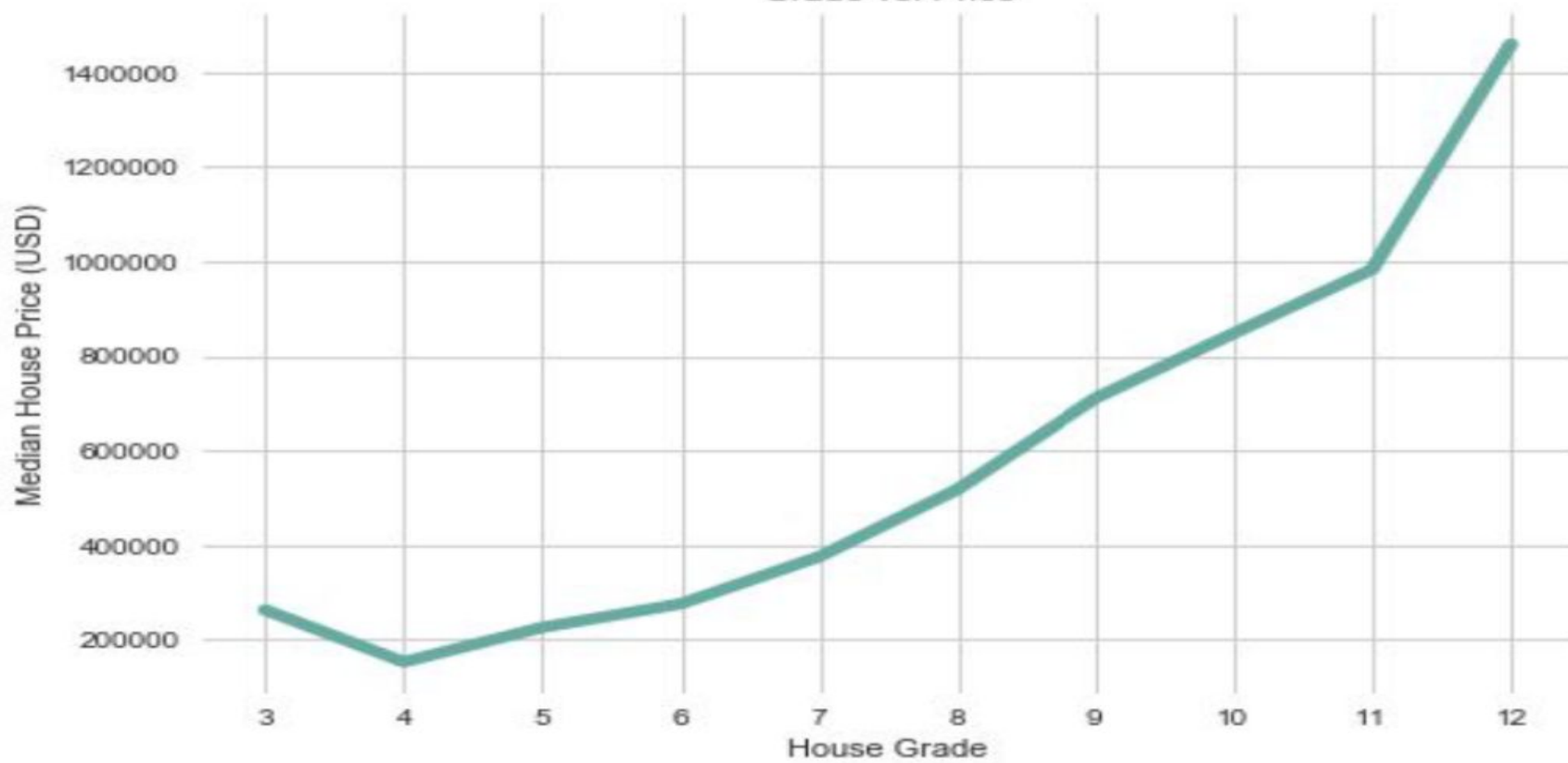
Results

Living Area Vs Price

Living area vs. Price
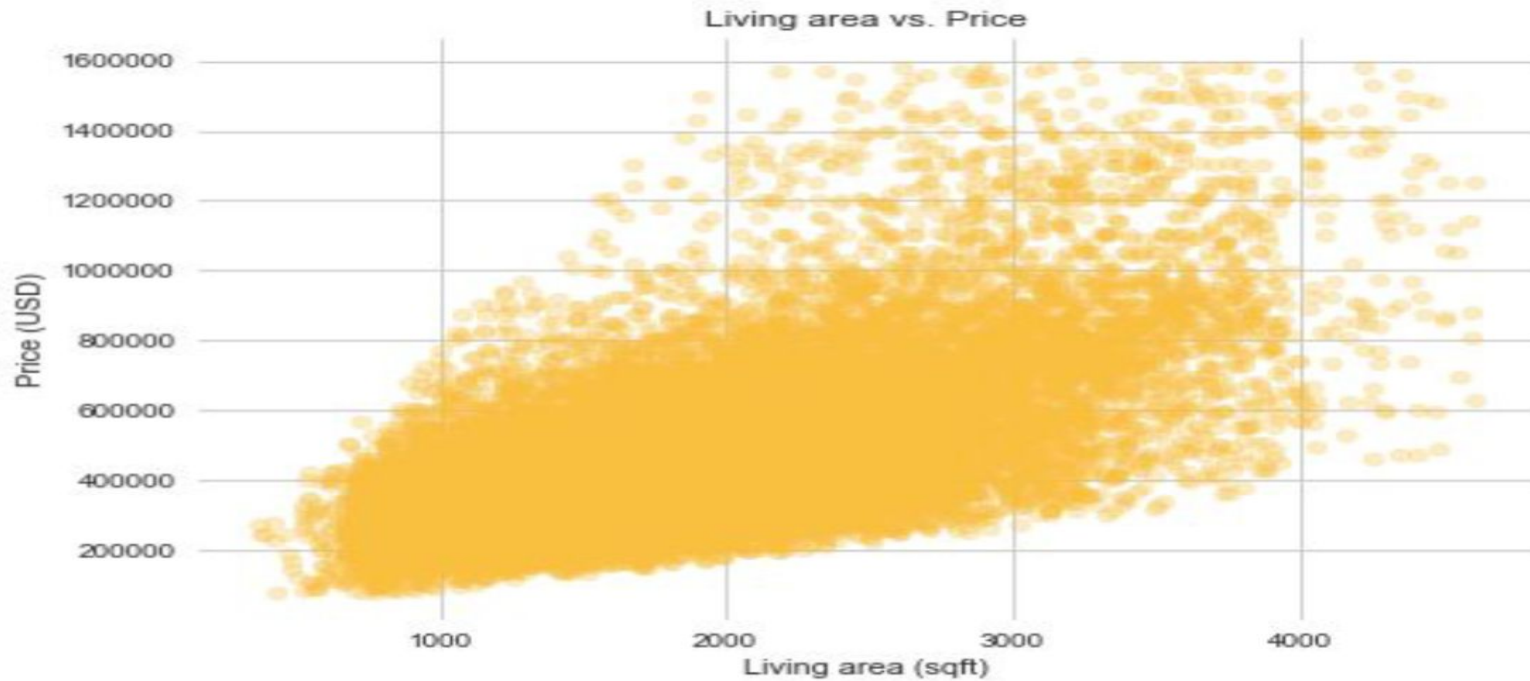
Living area vs. Price

Grade vs. Price

Recommendations

House Features

- Larger Living area(sqft)
- House grade higher than 10
- Grade is better than 10

Predicting House Prices

Let us now discuss our model for predicting house prices.

Predicting House Prices

	Model A(degree=2)	Model B(degree=3)
Features	12	12
Pros	Generalises well	Better Performance
Cons	Less accurate	
Mean Squared Error(MSE)	197491.97	196469.16
R-Squared	0.737	0.756

- Features is number of attributes used to predict.
- Score is a value between 0 and 1 with 1 being a perfect fit.
- Mean Squared Error is the amount in USD on average over or under.
- 73.7 % of variations in price can be explained by model A vs 75.6 % for model B

Conclusions

Model B is better at predicting house prices in Kings County by 75.6% compared to model A which is at 73.7%.

Thank You!

Email: mkanoru71@gmail.com
GitHub: @Kanoru01
LinkedIn: linkedin.com/in/KanoruMacharia/