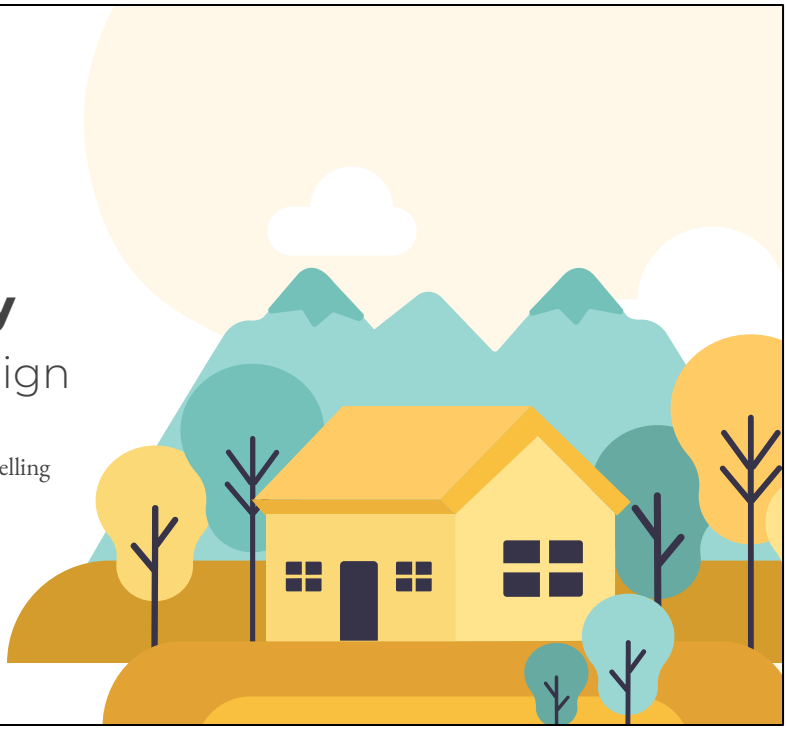
King County

Marketing Campaign

Data-driven insights and predictive modelling



Problem Statement

Advertising Campaign to encourage sales in King County

PropertiesInc

REAL ESTATE

Problem Statement

Advertising Campaign to encourage sales in King County

- Data-driven recommendations

PropertiesInc

REAL ESTATE

Problem Statement

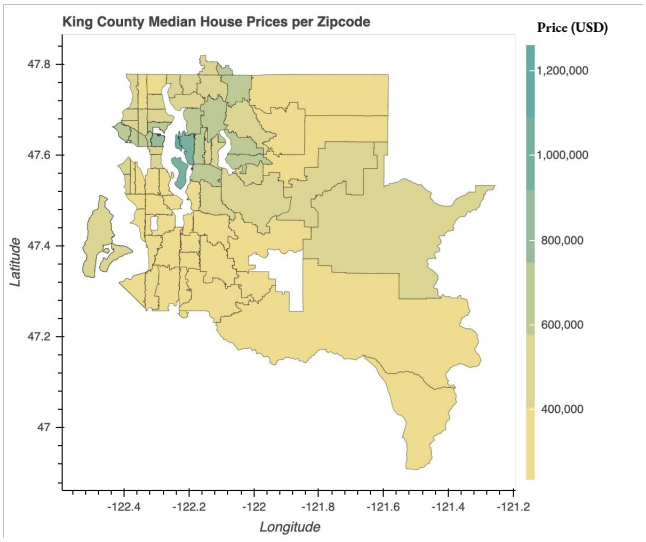
Advertising Campaign to encourage sales in King County

- Data-driven recommendations
- Model to predict house prices

Location

- Understand which areas have the highest house prices and should be the focus of our campaign

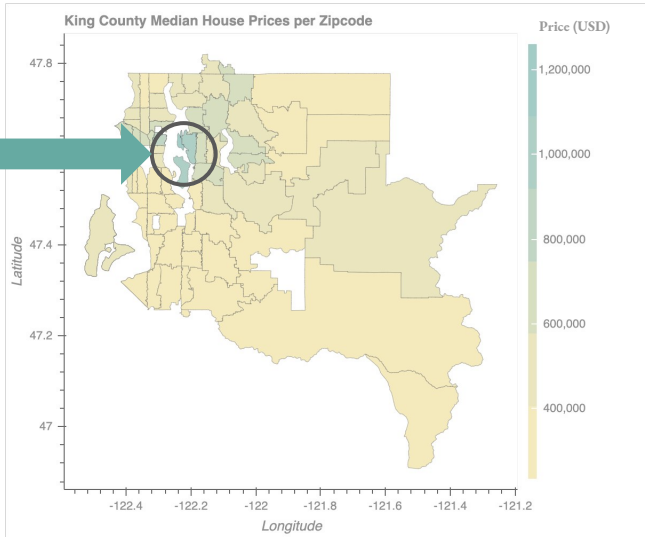
LOCATION - Zip Codes



- Looked at the median house price per zip code. There were 70 different zip codes in our dataset.

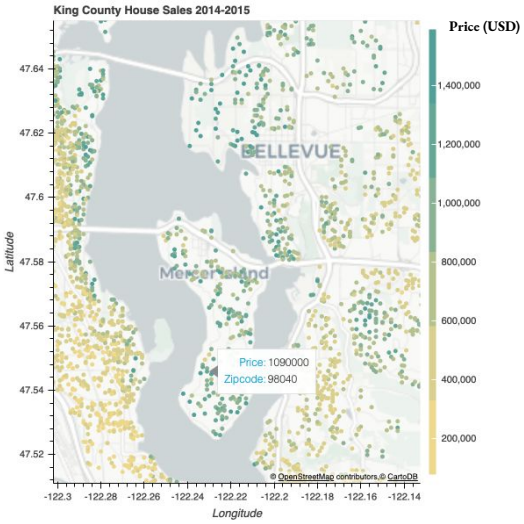
LOCATION - Zip Codes

Top 3 Zip Codes



- See top 3 zip codes in the NW area

LOCATION - Neighbourhoods



- Zoomed in and see that waterfront is key
- Most expensive houses are in Mercer Island and Bellevue

Recommendations

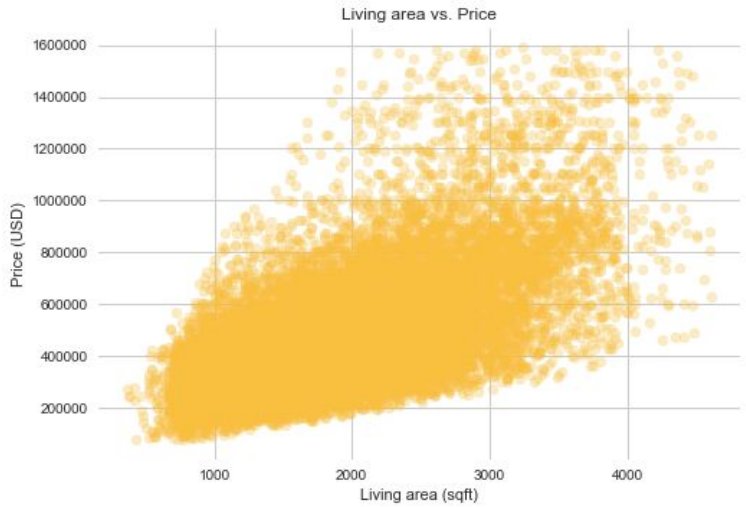
Location

- Importance of waterfront
- Bellevue and Mercer Island areas

House Features

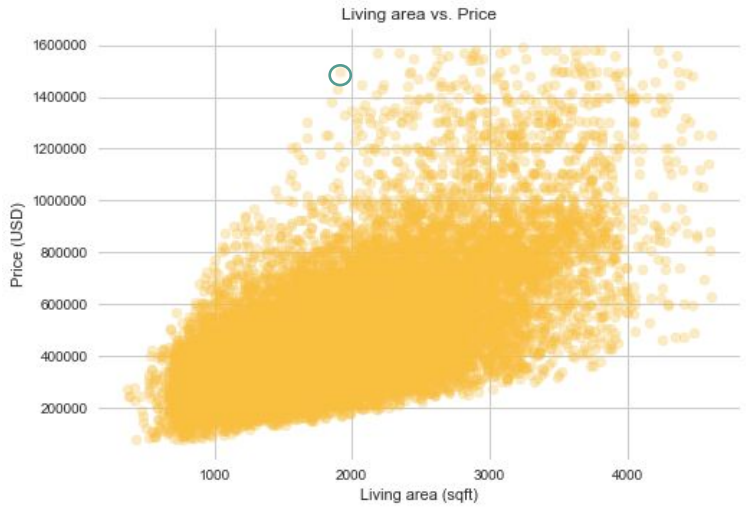
- Then looked at house features to see which factors drive the price up and also which factors are the best predictor of price

House Features - Living Area



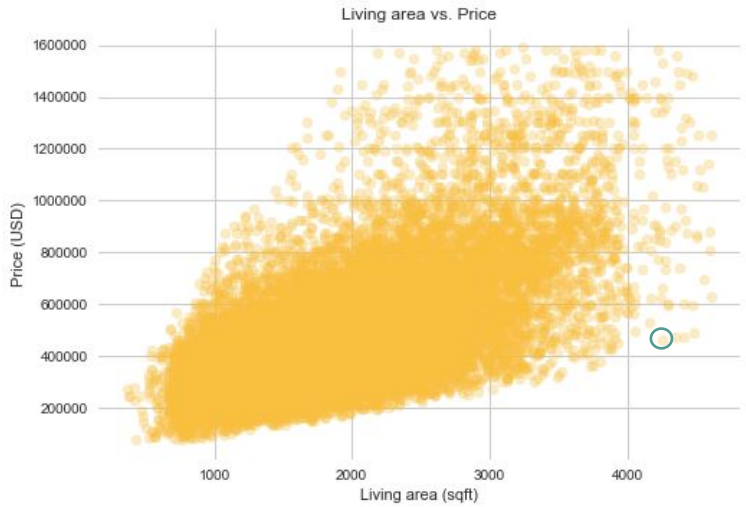
- Most obvious one first, sqft of living space. We would expect a larger house to be more expensive.
- See broad increase, more sqft is more expensive but not clear cut

House Features - Living Area



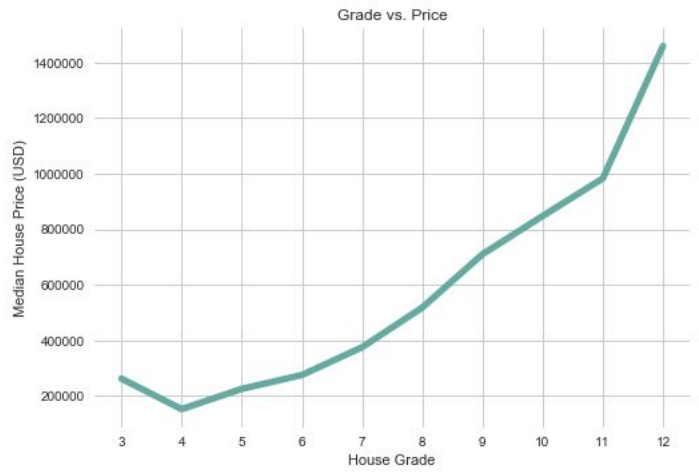
- Most obvious one first, sqft of living space. We would expect a larger house to be more expensive.
- See broad increase, more sqft is more expensive but not clear cut

House Features - Living Area



- Most obvious one first, sqft of living space. We would expect a larger house to be more expensive.
- See broad increase, more sqft is more expensive but not clear cut

House Features - Grade



- Better predictor is grade, construction value from King County which ranges from 1 to 13 (though here we only had 3 to 12)
- For the campaign we would recommend looking at houses with a grade of 10 or above. This would have a starting median price above $800,000. From the description we note that "Homes of this quality generally have high quality features. Finish work is better and more design quality is seen in the floor plans. Generally have a larger square footage."

House Features - Other

Was viewed

Presence of Basement

Square Feet of Lot

Recently renovated

Condition

Bedroom count

Year Built

Waterfront view

Bathroom count

Number of floors

- Other features we considered

Recommendations

Home Features

- Larger living area (sqft)
- Grade 10 and above
- Grade better predictor

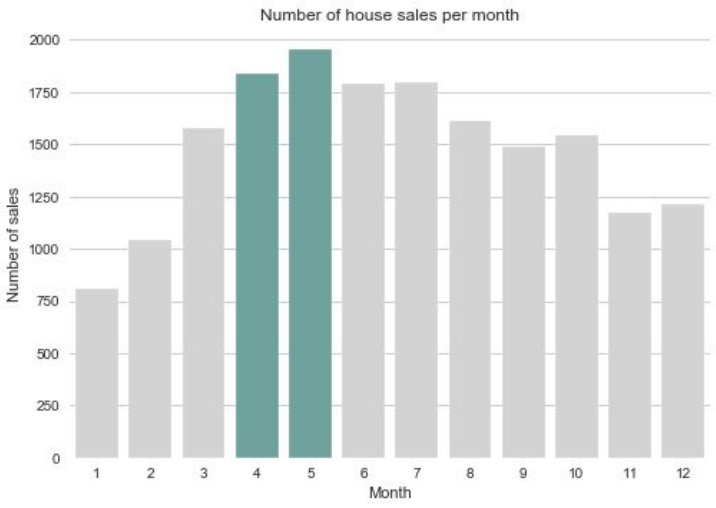
Location

- Importance of waterfront
- Bellevue and Mercer Island areas

Time of Year

- Investigated whether the median price was higher at a certain time of the year and whether there are any trends.
- Median price was almost the same, so no influence

Time of year



- However volume of sales varied
- April and May are the most popular months for house sales. In contrast, January and February have the lowest number of sales.
- Recommend March/April for the campaign

Recommendations

Home Features

- Larger living area (sqft)
- Grade above 10
- Grade better predictor

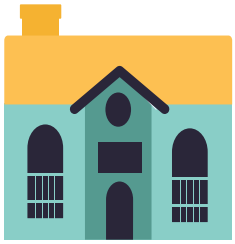
Location

- Importance of waterfront
- Bellevue and Mercer Island areas

Time of Year

- Constant median house price
- April/May highest sales volume
- Campaign in March/April

Predicting House Prices



Predicting House Prices

Features		
Pros		
Cons		
Score		
Mean Error		

- Features is number of attributes used to predict
- Score is a value between 0 and 1 with 1 being a perfect fit
- Mean error is the amount in USD on average over or under

Predicting House Prices

	Model A	
Features	17	
Pros		
Cons		
Score		
Mean Error		

Predicting House Prices

	Model A	
Features	17	
Pros	Easy to interpret, generalises well	
Cons		
Score		
Mean Error		

Predicting House Prices

	Model A	
Features	17	
Pros	Easy to interpret, generalises well	
Cons	Less accurate	
Score		
Mean Error		

Predicting House Prices

	Model A	
Features	17	
Pros	Easy to interpret, generalises well	
Cons	Less accurate	
Score	0.70	
Mean Error		

- 70% of variations in price can be explained by model A vs 83% for model B

Predicting House Prices

	Model A	
Features	17	
Pros	Easy to interpret, generalises well	
Cons	Less accurate	
Score	0.70	
Mean Error	USD 132,444	

Predicting House Prices

	Model A	Model B
Features	17	87
Pros	Easy to interpret, generalises well	
Cons	Less accurate	
Score	0.70	
Mean Error	USD 132,444	

Predicting House Prices

	Model A	Model B
Features	17	87
Pros	Easy to interpret, generalises well	Performance
Cons	Less accurate	
Score	0.70	
Mean Error	USD 132,444	

Predicting House Prices

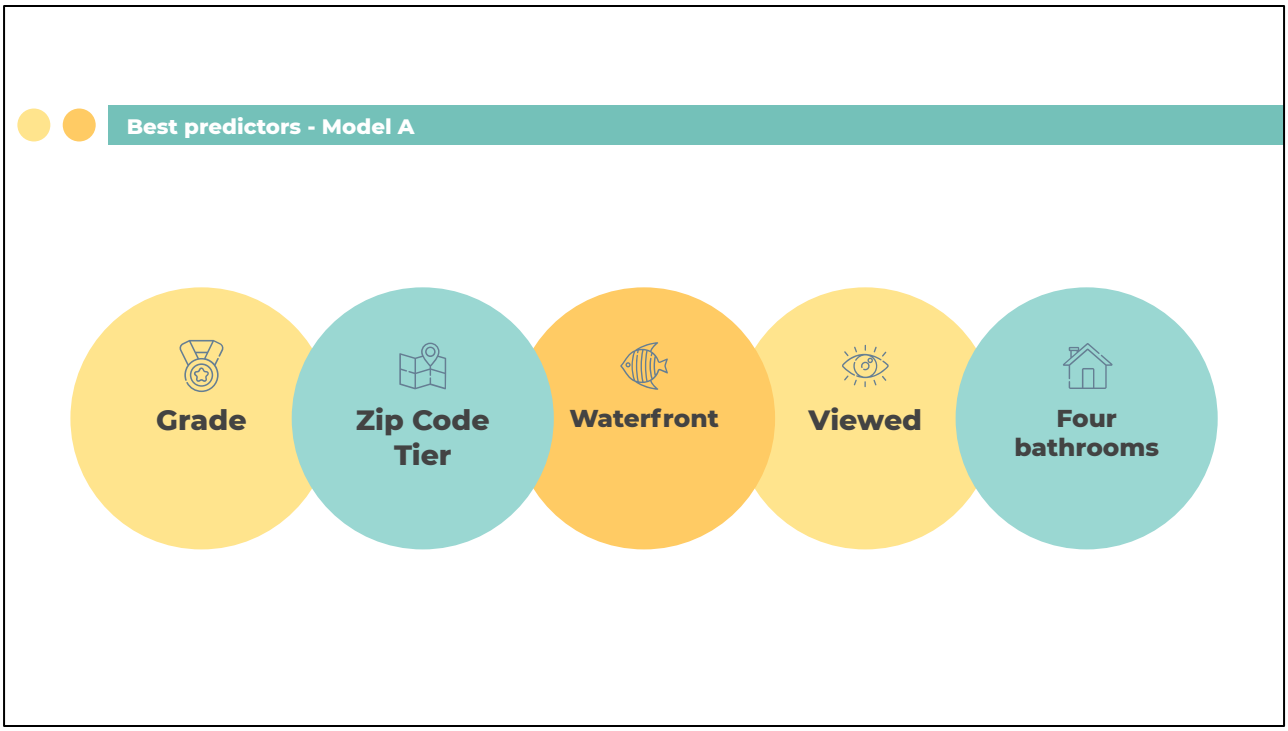
	Model A	Model B
Features	17	87
Pros	Easy to interpret, generalises well	Performance
Cons	Less accurate	Uses exact zip code data
Score	0.70	
Mean Error	USD 132,444	

Predicting House Prices

	Model A	Model B
Features	17	87
Pros	Easy to interpret, generalises well	Performance
Cons	Less accurate	Uses exact zip code data
Score	0.70	0.83
Mean Error	USD 132,444	

Predicting House Prices

	Model A	Model B
Features	17	87
Pros	Easy to interpret, generalises well	Performance
Cons	Less accurate	Uses exact zip code data
Score	0.70	0.83
Mean Error	USD 132,444	USD 99,654



- Grade 12 is worth $52,000 more than grade 11
- being on the waterfront is valued at USD 277,442.

Further work

School Ratings Data	Proximity to a good school is likely to increase sale price

Further work

School Ratings Data	Proximity to a good school is likely to increase sale price
Commuting Time Data	Homes better connected to downtown Seattle are likely to be more valuable

Further work

School Ratings Data	Proximity to a good school is likely to increase sale price
Commuting Time Data	Homes better connected to downtown Seattle are likely to be more valuable
Longer Time Scale	See which areas show signs of growth/ decline

THANK YOU

Nadine Amersi-Belton
nzamersi@gmail.com
https://github.com/nadinezab

APPENDIX - Model A Coefficients

	Feature	Coefficient		Feature	Coefficient
1	grade_12	515,536	10	zip_tier_5	168,374
2	grade_11	463,199	11	viewed	153,790
3	zip_tier_1	435,924	12	zip_tier_6	133,920
4	grade_10	327,375	13	bathrooms_4	105,063
5	zip_tier_2	313,144	14	zip_tier_7	99,454
6	waterfront	277,442	15	grade_8	88,713
7	zip_tier_3	250,036	16	grade_5	-94,930
8	zip_tier_4	248,289	17	grade_4	-136,094
9	grade_9	233,634