Predicting Housing Prices in King County, WA

Hazel Donaldson

Introduction

Data: King County Housing Sales

Goal: Create a model to predict future prices

Will briefly summarize methodology and some of my findings

Overview of Methodology OSEMiN

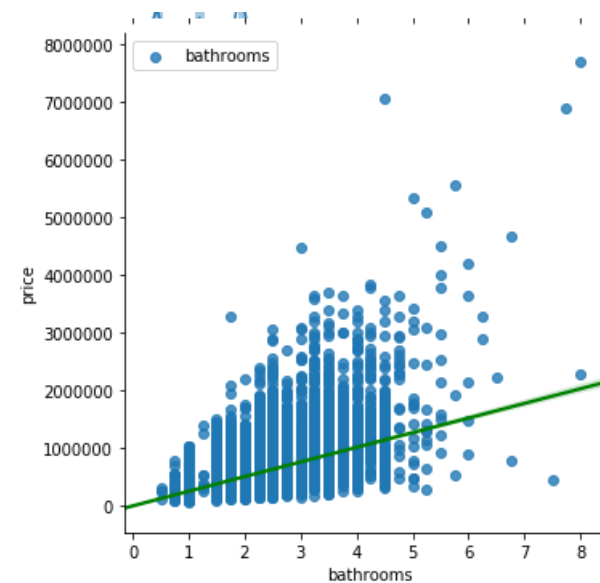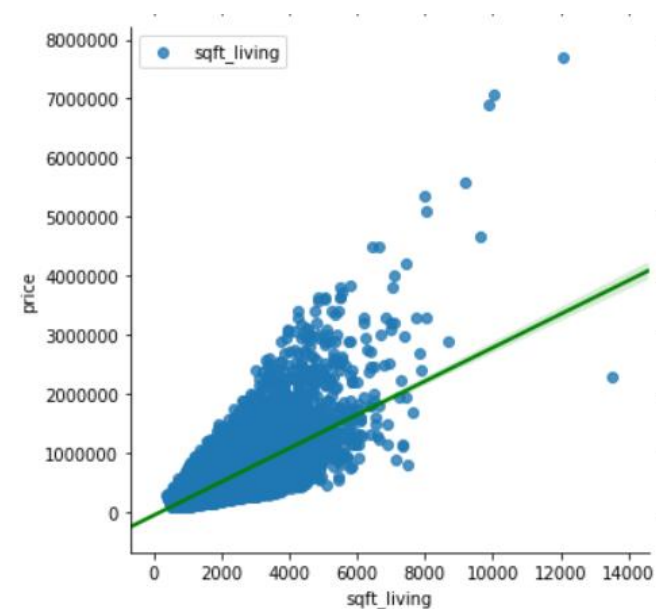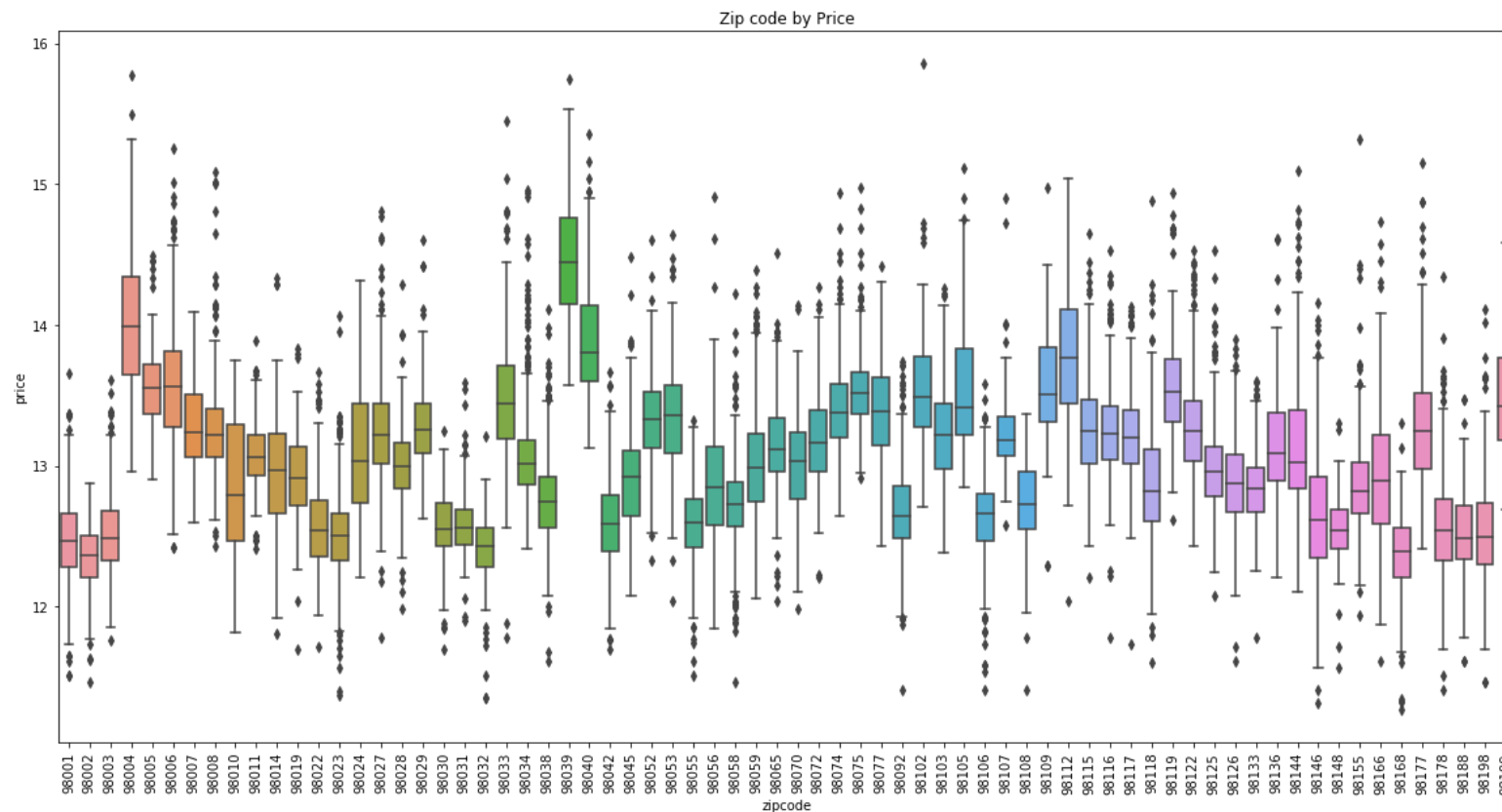|		
Obtain	Obtain the Data
Scrub	Scrub the Data
Explore	Explore the Data
Model	Model the Data
Interpret	Interpret the Data

Prediction Features

- Bedrooms
- Bathrooms
- Floors
- Condition
- View
- Waterfront
- Grade
- Sale Price

- Year Built
- Date of Sale
- Year Renovated
- Sqft living
- Sqft basement
- Sqft lot
- Zip code
- Longitude/latitude

How Features affect Price
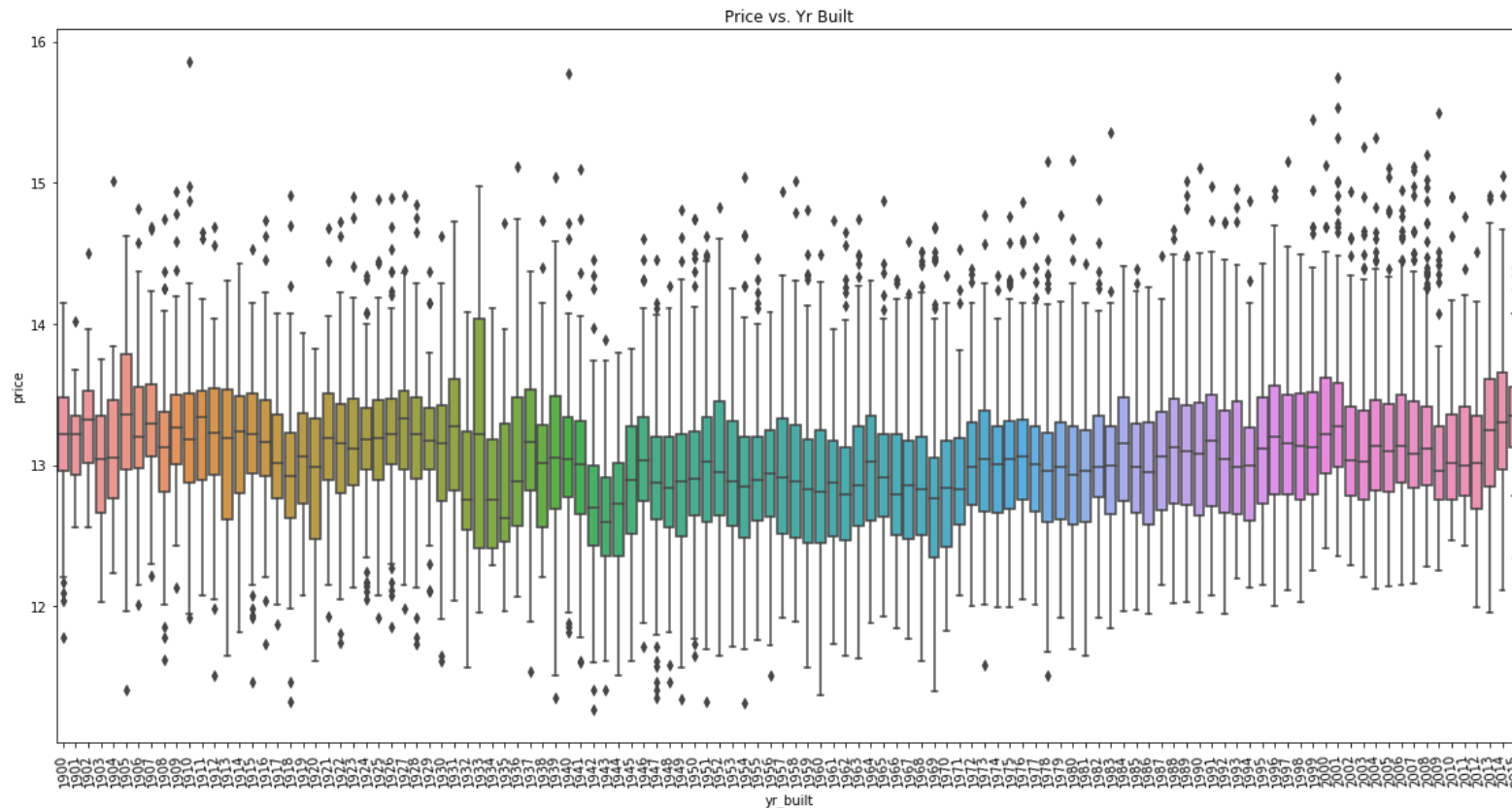
Does location affect price?

Zip code by Price

Location vs. Price cont.

Most expensive zip codes:

- 98004: Medina
- 98039: Bellevue
- 98040: Mercer Island
- 98112: N.E. Seattle
- 98199: N.W. Seattle

Are newer homes more expensive?

Price vs. Yr Built

- No clear increase in price

Selected Features

- Location(zip code)
- Sqft_living
- Grade
- Bathrooms
- View

Dep. Variable:	price	R-squared:	0.881
Model:	OLS	Adj. R-squared:	0.880
Method:	Least Squares	F-statistic:	1762.
Date:	Tue, 18 Jun 2019	Prob (F-statistic):	0.00
Time:	23:13:01	Log-Likelihood:	5454.7
No. Observations:	19163	AIC:	-1.075e+04
Df Residuals:	19082	BIC:	-1.011e+04
Df Model:	80		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	7.8858	0.050	156.542	0.000	7.787	7.985
waterfront[T.1.0]	0.4993	0.017	29.420	0.000	0.466	0.533
bedrooms	-0.0134	0.002	-6.685	0.000	-0.017	-0.009
bathrooms	0.0331	0.003	11.135	0.000	0.027	0.039
sqft_living	0.3747	0.007	52.139	0.000	0.361	0.389
sqft_lot	0.0814	0.004	21.332	0.000	0.074	0.089
floors	0.0127	0.003	3.811	0.000	0.006	0.019
view	0.0579	0.002	27.799	0.000	0.054	0.062

Recommendations

- Square footage of living space is important
- High grade buildings
- Neighborhoods with large homes
- Focus on top 5 zip codes (northwest of the county is best)

Future Work

- To improve model:
  - Does date of sale affect price?
  - How does renovation affect price?
  - Why number of bedrooms didn't have positive impact on price?

Thank you for your time!