PREDICTING HOME VALUES

Jamaal Smith

Table of Contents

- **Executive Summary**

- Importance of Home Values in the United States

- Key Predictor Variables

- Discussion of Regression Outputs

- Key TakeAways

Executive Summary

- Houses represent the largest asset that most Americans will purchase in their lifetimes and is often the key driver of a family's net worth.

- Objective measures of predicting home values has multiple applications in the real-world.

- This analysis can help communities better understand whether other drivers such as proximity to employment drive the appreciation of home prices more than a house's intrinsic features such as number of bedrooms

Table of Contents

- Executive Summary

- **Importance of Home Values in the United States**

- Key Predictor Variables
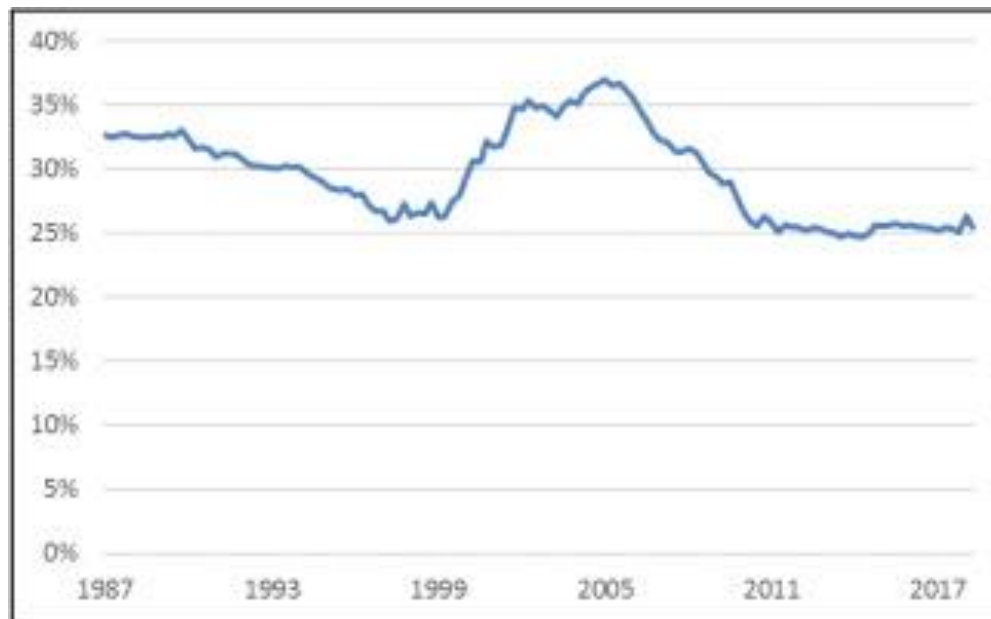
- Discussion of Regression Outputs

- Recap

Home Values Versus Stock Market Returns

Home Value ($000)

Stock Market Returns

Source: US Census Bureau & Dow Jones
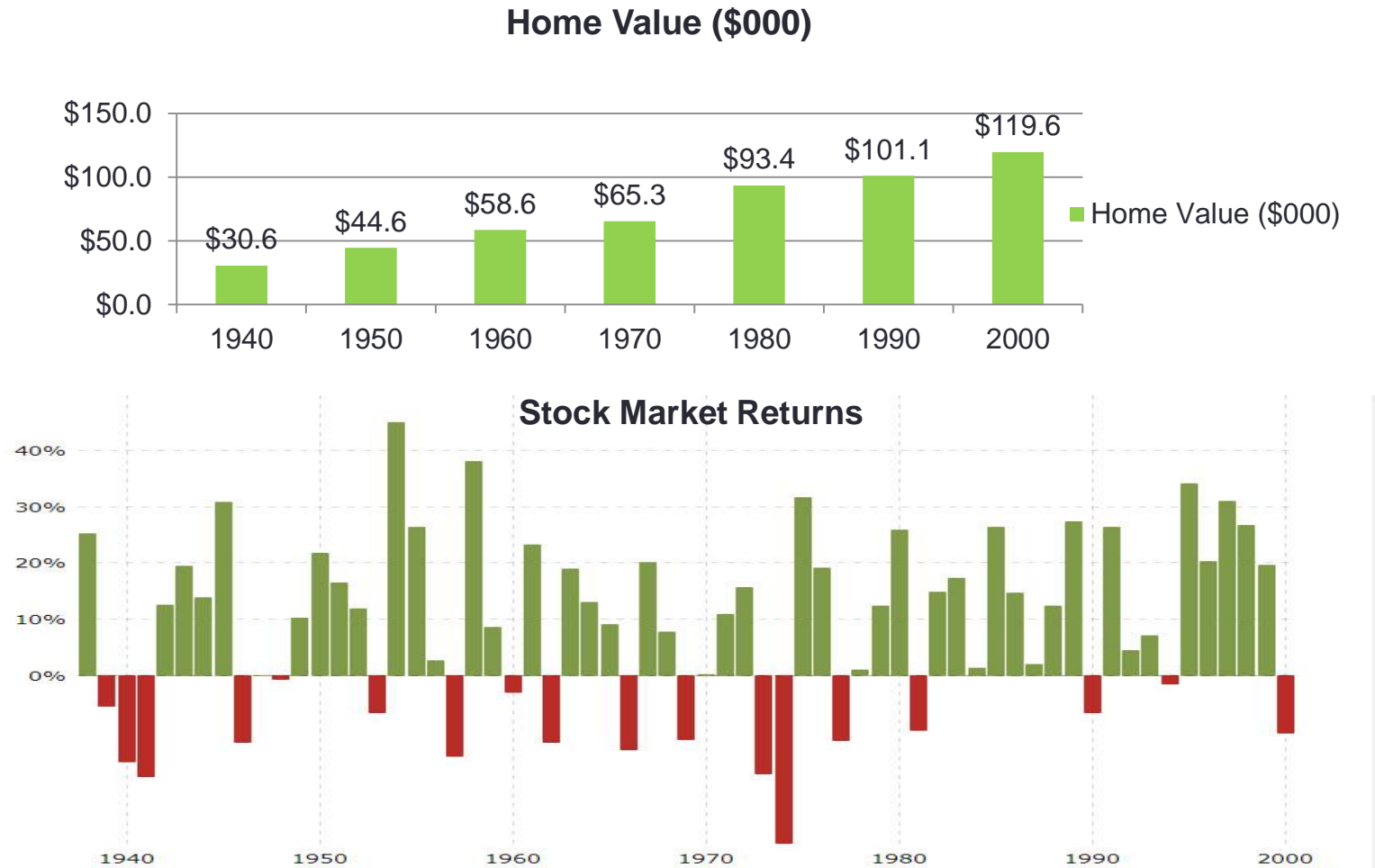
Housing Values Key To US Economy

- A house is often the only major asset American families acquire and drives their net worth.

% of American Household Wealth in Home

Declines in Housing Prices Often Coincide with Economic Decline

- In recent history, steep declines in home values have resulted in economic hardship in the United States

Year-over-year Housing Price Changes

Source: FHFA, House Price Indexes, Seasonally Adjusted Purchase-Only Index

Sales of Existing Homes Represent Majority of Housing Sales

- With existing homes representing the majority of housing sales, it is important to develop objective manners to predict the sale price of existing homes
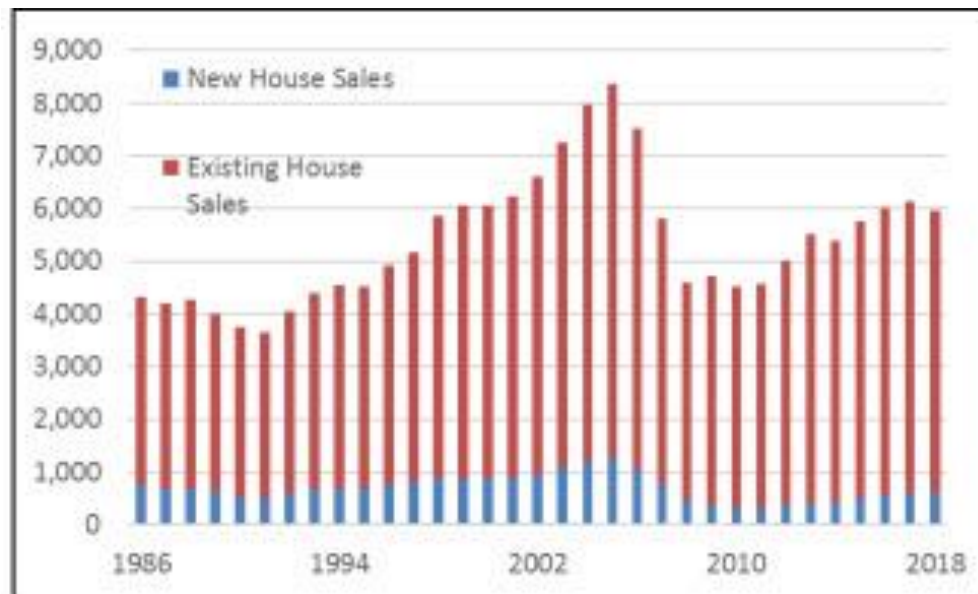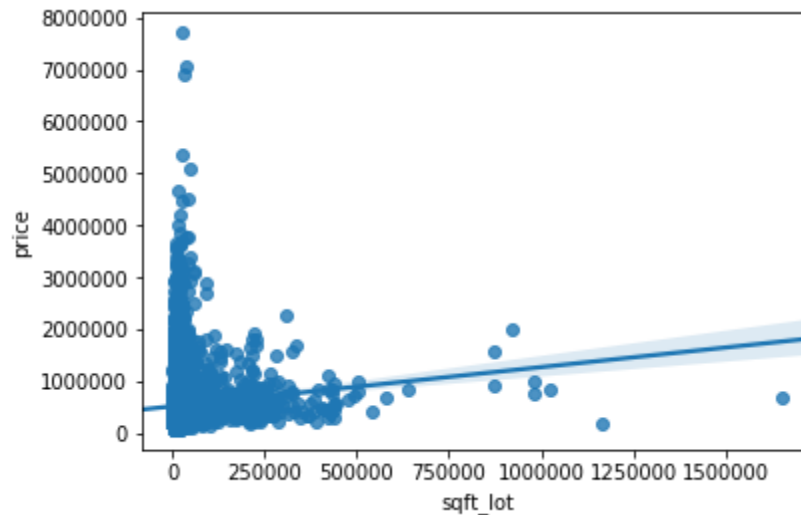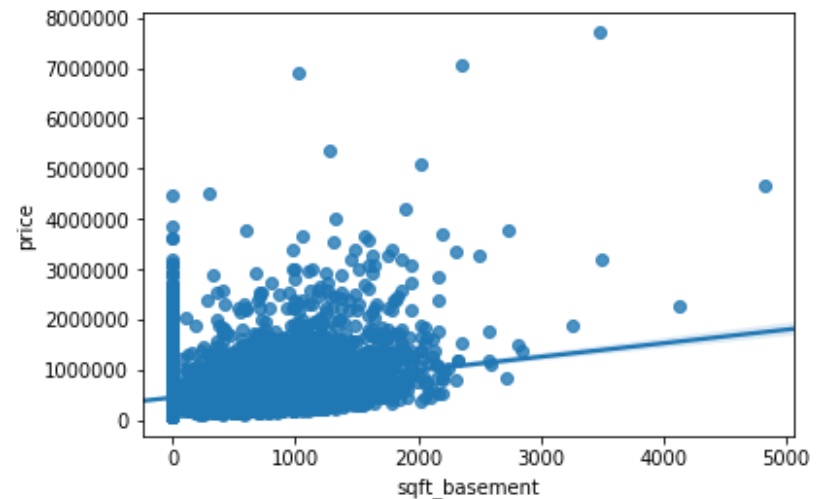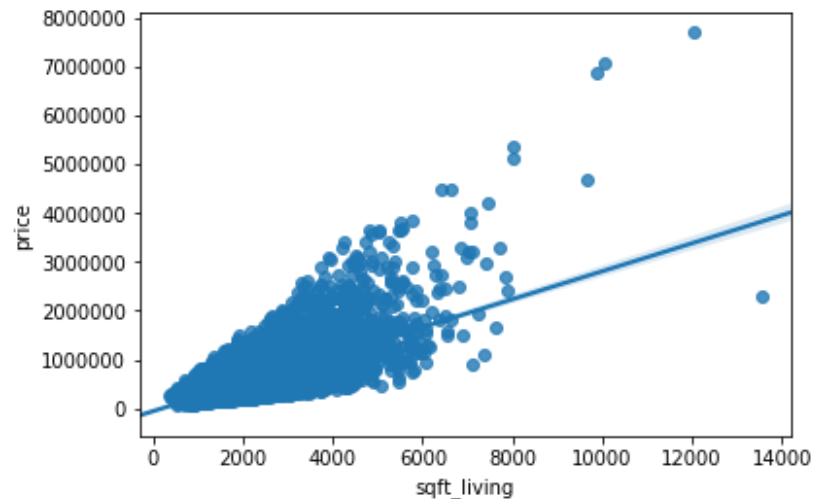
Table of Contents

- Executive Summary

- Importance of Home Values in the United States

- **Key Predictor Variables**

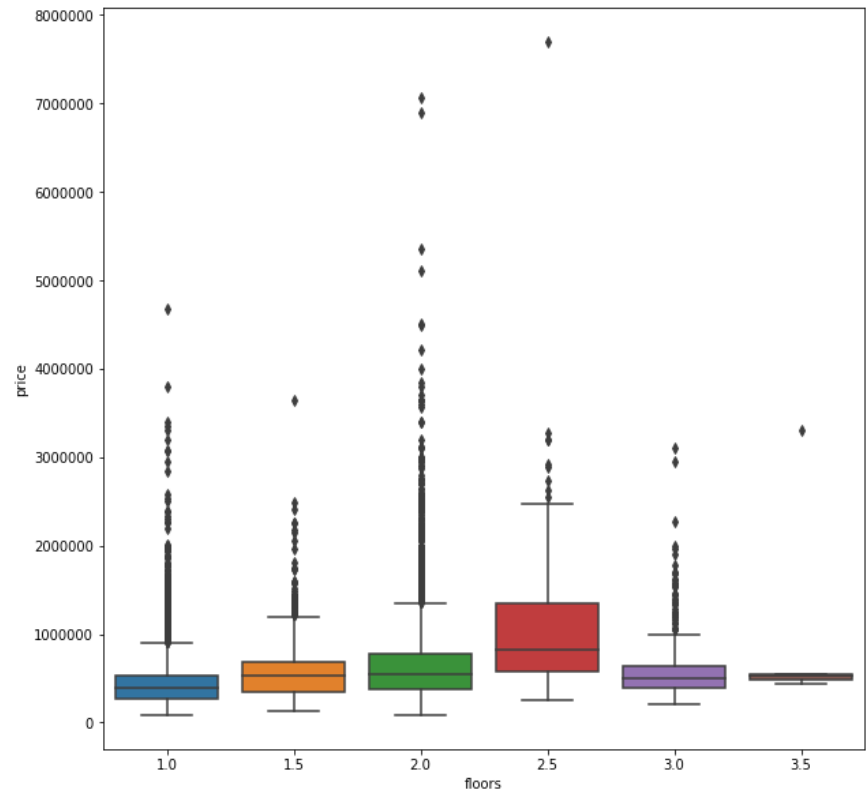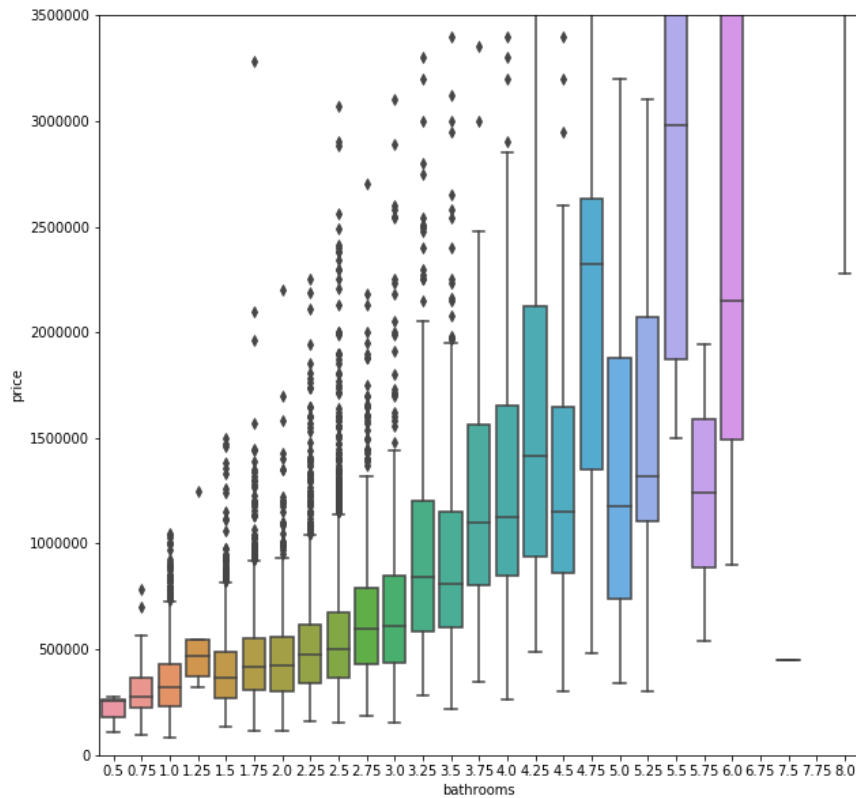- Discussion of Regression Outputs

- Reca[

Basic Data Set Details

- Data Set analyzed contained sales price and housing characteristics of homes sold in the Seattle area between May 2014 and May 2015.

- Each house sold had roughly 20 data points that are represented as columns in the spreadsheet
  - For this analysis, the column headers are the predictor variables that can possibly be included in the regression model

- The relationship between a home's sale price and a specific characteristic is important to consider when determining whether or not to include a predictor variable

Predictor Variable Relationship Type

Predictor Variable Relationship To Price
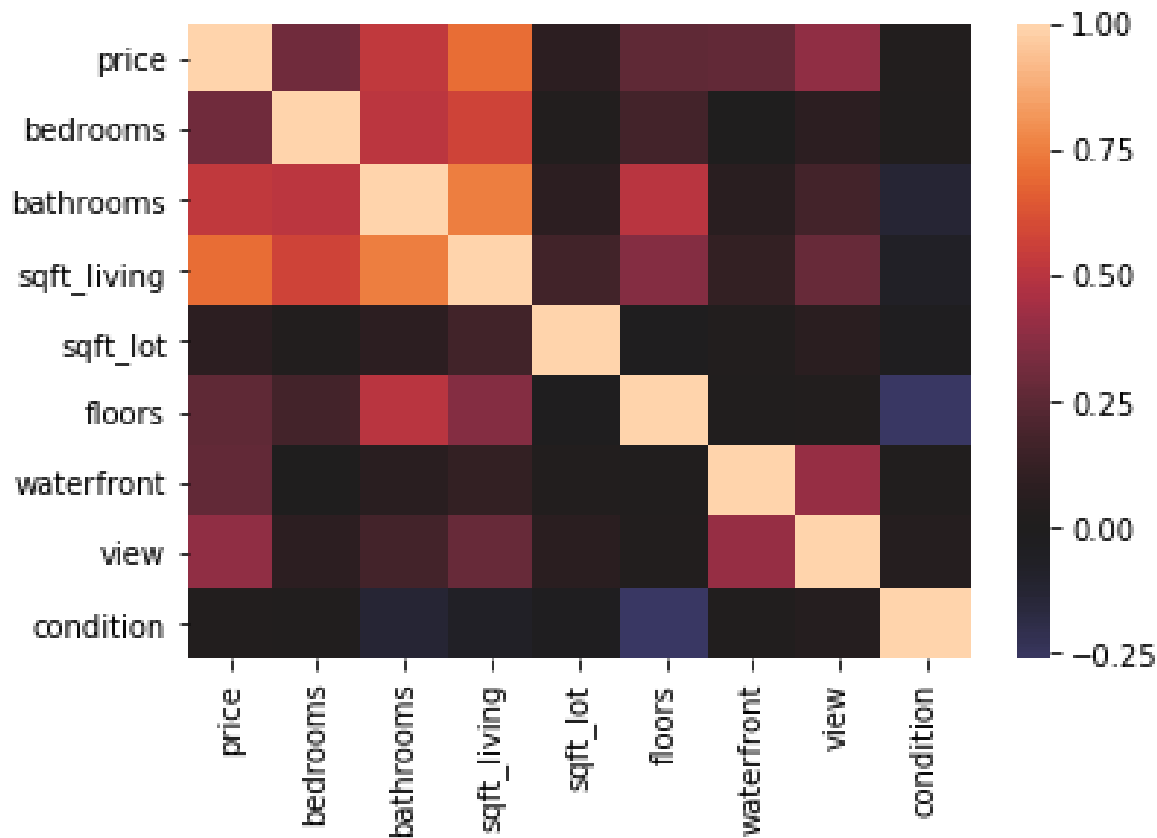
Multicollinearity treatment

Table of Contents

- Executive Summary

- Importance of Home Values in the United States

- Key Predictor Variables

- **Discussion of Regression Outputs**

- Recap

Rating the Model

OLS Regression Results

Dep. Variable:	price	R-squared:	0.771
Model:	OLS	Adj. R-squared:	0.770
Method:	Least Squares	F-statistic:	3451.
Date:	Tue, 19 Nov 2019	Prob (F-statistic):	0.00
Time:	00:15:27	Log-Likelihood:	-640.33
No. Observations:	15429	AIC:	1313.
Df Residuals:	15413	BIC:	1435.
Df Model:	15		
Covariance Type:	nonrobust		

- Our linear regression model can predict the sale price of a home in Kings County with 77% accuracy.
- The results of this regression are statistically significant as the probability of the F-statistic is below 5%.
- With additional measures, the model can be improved to be more accurate and take into consideration the location of a home

Variables that Impact Price

```
price              1.000000
sqft_living        0.706498
grade              0.664356
sqft_above         0.612374
sqft_living15      0.583400
bathrooms          0.525655
view               0.397786
sqft_basement      0.320053
lat                0.306991
bedrooms           0.305792
waterfront         0.275718
floors             0.260143
yr_renovated       0.120981
sqft_lot           0.084687
sqft_lot15         0.080569
yr_built           0.050528
condition          0.033172
long               0.021948
date              -0.000445
id                -0.014954
zipcode           -0.049371
```

- The amount of square feet in a home's living space is the highest correlated variable to its sales price.

- Variables such as date, id and zipcode might be excluded from the final model as they have negative correlations with a home's sales price

Table of Contents

- Executive Summary

- Importance of Home Values in the United States

- Key Predictor Variables

- Discussion of Regression Outputs

- **Recap**

Important Points

- The linear relationship between the predictor variables and sale price of homes and presence of low multicollinearity met two of the assumptions for linear regression. Attempts were also made to normalize the data and residuals for the model were not skewed.

- Some key predictors of a home's sale price are the square feet of living space, its number of bathrooms and relative size of other homes in the area as shown in the previous slide.

- The ability to minimize the model's error may not be ideal as market factors beyond a home's features play a role in a home's final sale price.