Zillow Regression

By Chris Everts

The Hottest trend

Agenda

- Executive Summary
- The Source
- What is Driving Prices
- The Crystal Ball
- Taxes in California
- Conclusion

Executive Summary

- Data for single unit properties from May-August, 2017
- Key drivers of home value
  - Square Feet
  - Bedrooms
  - Bathrooms
- Polynomial Regression = Best model for predicting value
- Tax Distribution by rank
  - Las Angeles
  - Orange
  - Ventura

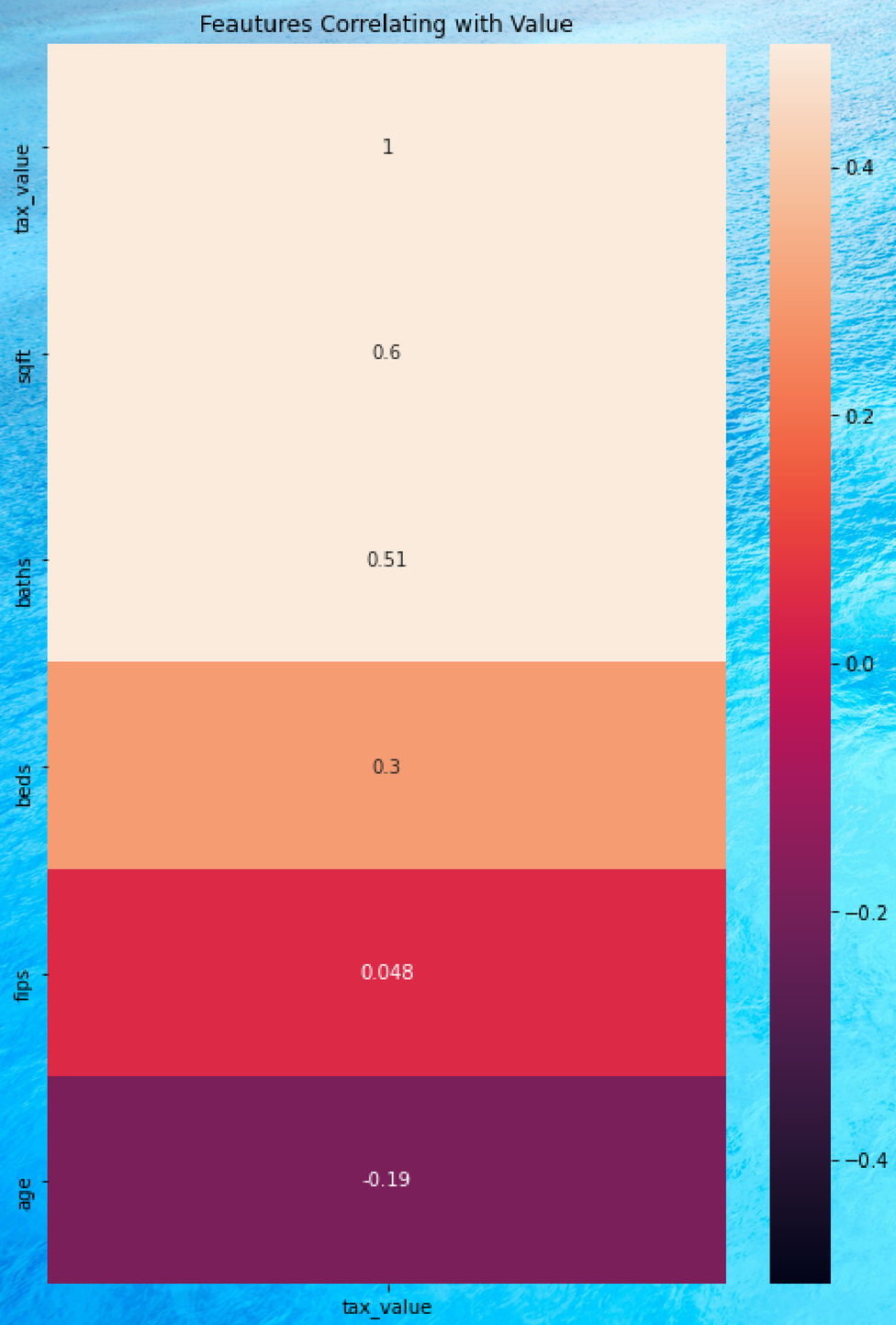
The Source

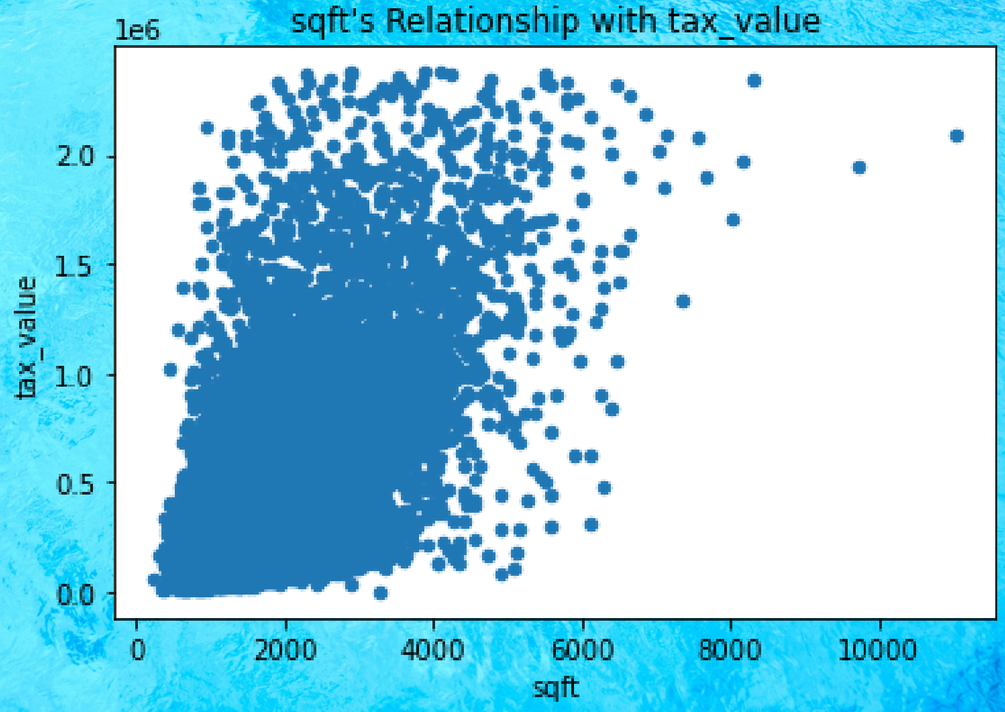
- Single unit property records from May-August, 2017
- Location: California
- Condensed data down to 38,000 rows
- Max sq. ft of 11,000 and average of 1700 sq. ft.

The Drive

The Top 3 Drivers of Value

1. # of Square Feet
2. # of Bathrooms
3. # of Bedrooms





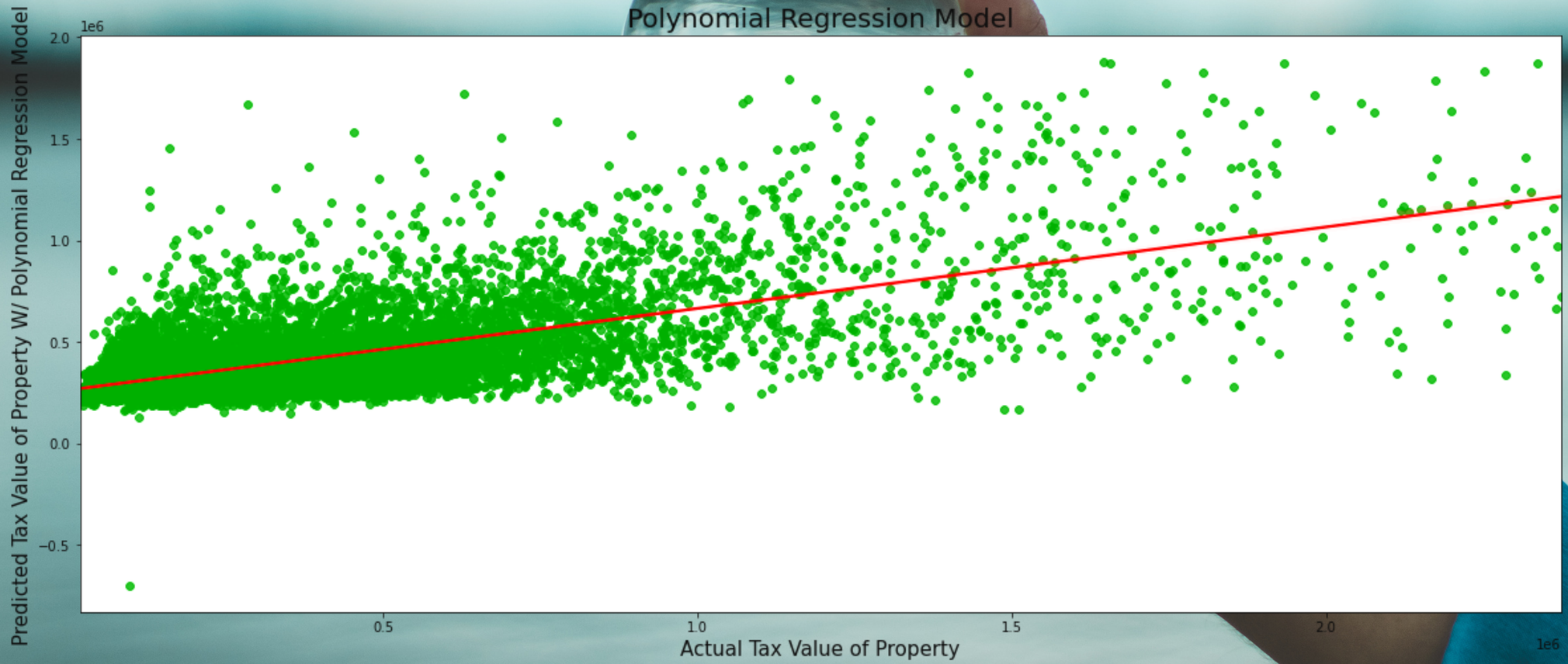
The Crystal Model

Baseline

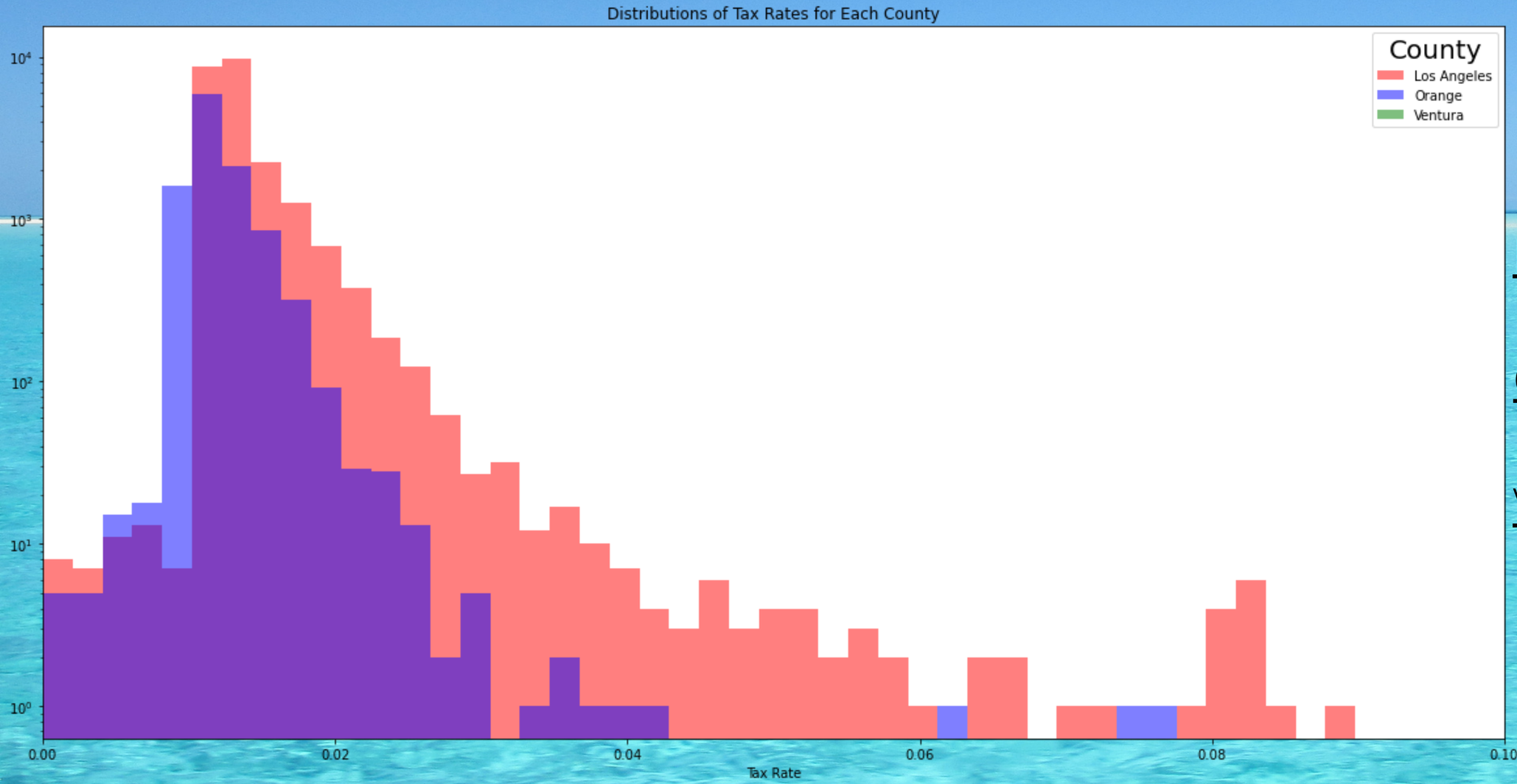
RMSE using Mean
Train/In-Sample: 357,185.61$
Validate/Out-of-Sample: 359,454.06$
The R squared score for the baseline is -3.0

Polynomial Regression Model

RMSE for Polynomial Model, degrees=3
Training/In-Sample: 278,845.48$
Validation/Out-of-Sample: 277,370.689$
The r^2 value for Polynomial is 0.4



Tax Distributions



Tax Rate

Los Angeles 1.38%

Orange 1.20%

Ventura 1.19%

Conclusion

Drivers of Value

Square Foot

Bedrooms

Bathrooms

The Prediction

Polynomial Model performed best

Tax Rate

Highest tax rate was found in LA

Next Steps

Add more features to run more models

Ask for data on out of state buyers for further analysis