

# HOUSING PRICE PREDICTION

...

An Nguyen

Advisors: Chris Fernandes, Nick Webb & Harlan Holt

# 0. BACKGROUND

**Zillow**

**34 Shadow Wood Ln,  
Thomaston, GA 30286**

**3 beds · 2 baths · 1,805 sqft**

**SOLD: \$160,000**  
Sold on 01/12/18

**Zestimate®:**  
**\$184,777**

**EST. REFI PAYMENT**  
**\$645/mo**

[See current rates](#)

**Home Shoppers  
are Waiting**

Ask an agent about market conditions in your neighborhood.

Your name  
Phone  
Email

Features:

Beds  
Baths  
Size

Sold Price:  
\$160,000

Zestimate:  
\$184,777

# 0. BACKGROUND



- 26.2% market share
- 110 million houses
- Zestimate



- 9.6% market share



- 3.5% market share

# 1. PROBLEMS

- Zillow correctly estimates ~50% of their houses within 5% range of the actual sold price

## 2. QUESTIONS

- Can I get close to or beat the Zestimate?

# 1. PROBLEMS

- Zillow tends to **overestimate** their properties

## 2. QUESTIONS

- Can my models get rid of the overestimation problem?

# 1. PROBLEMS

- Do we need a lot of **attributes** to have a good prediction for house price?



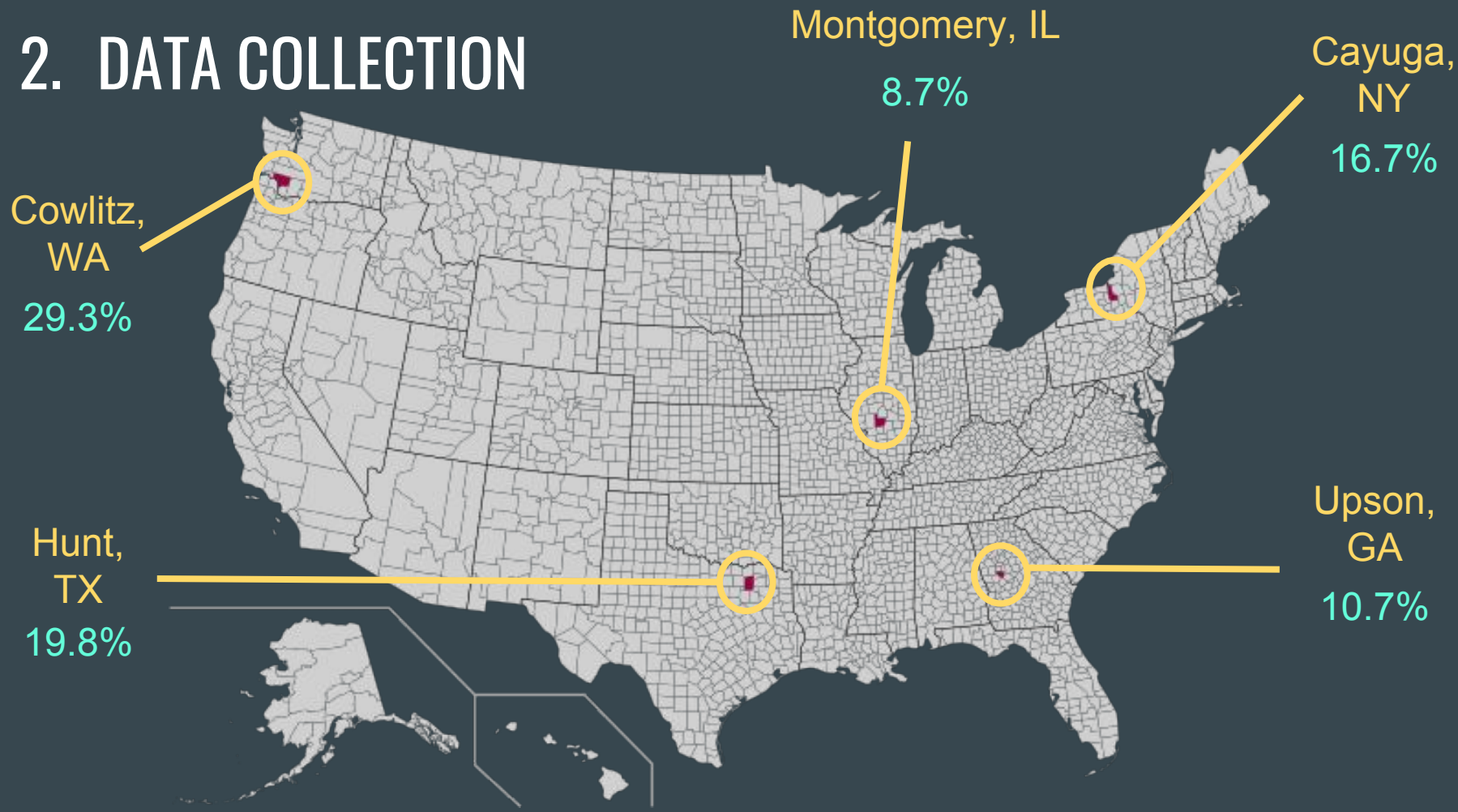
## 2. QUESTIONS

- Most important attributes?

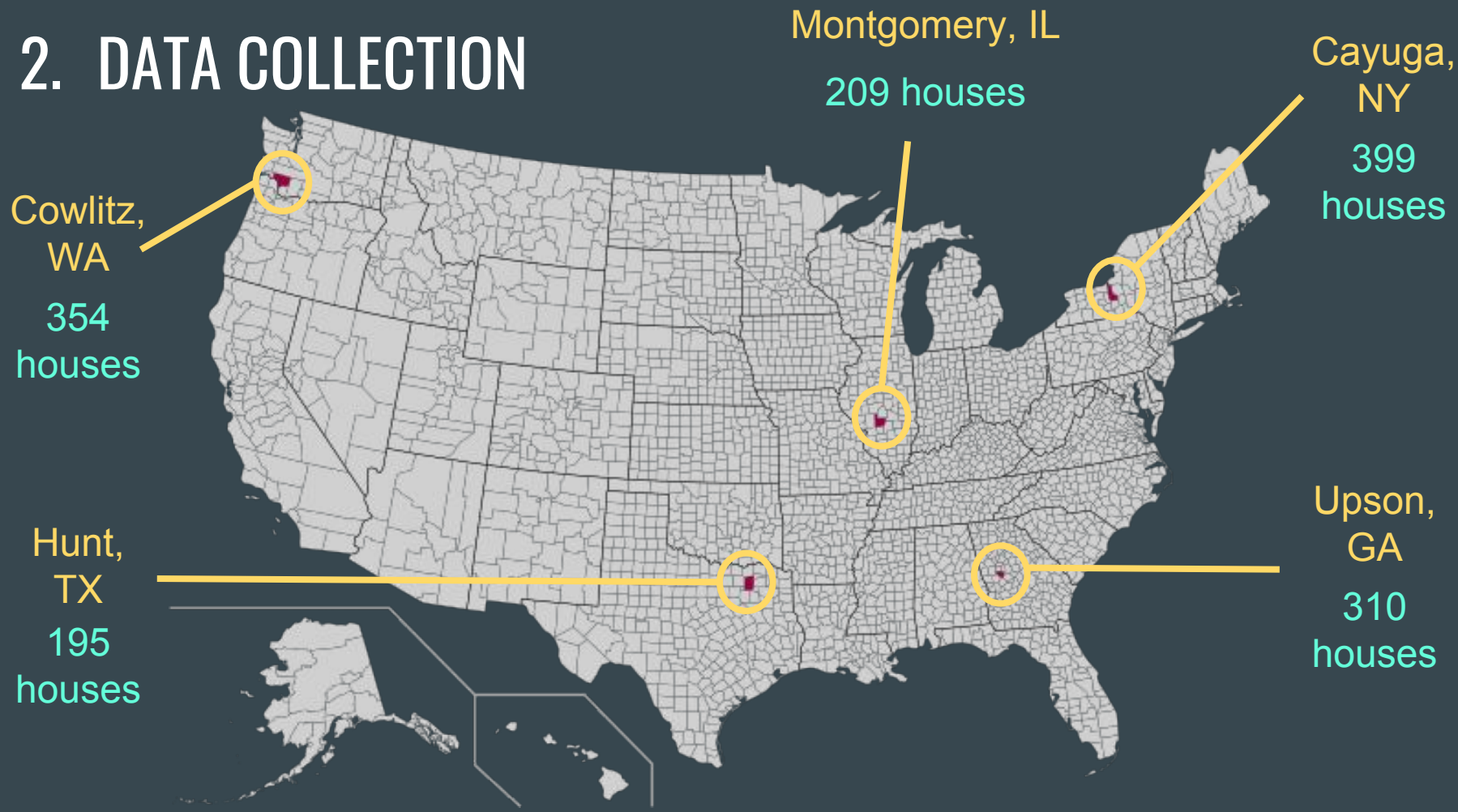
## 2. QUESTIONS

- Can I get close to or beat the Zestimate?
- Can my models get rid of the overestimation problem?
- Most important attributes?

## 2. DATA COLLECTION



## 2. DATA COLLECTION



## 2. DATA COLLECTION

- Sources: Zillow, Trulia, and Redfin
- Tools: Python, Selenium, and VBA
- Attributes:
  - Internal Factors: Beds, Baths, Size, Appliances, Garage, etc.
  - External Factors: Tax Info, School Info, Walkability, Nearby Lifestyle Amenities, Comparable Houses' Sold Prices

**215 P G Sweet Rd**  
Kelso, WA 98626  
Status: Sold

**\$313,342**  
Redfin Estimate

**\$325,000**  
Last Sold Price

**4** Beds  
**1.75** Baths

Overview Property Details Tour Insights Redfin Estimate Property History Public F



**215 P G Sweet Rd**  
Kelso, WA 98626  
3 beds • 1.5 baths • 1,375 sqft • 0.7 acres lot size • Single-Family Home

OFF MARKET • **SOLD** on Oct 3, 2017  
**\$233,427**



**215 P G Sweet Rd,**  
**Kelso, WA 98626**

**4 beds • 2 baths • 2,275 sqft**

**SOLD: \$325,000**  
Sold on 10/03/17  
Zestimate®:  
\$240,487

Zillow

\$233,427

Trulia

\$325,000

\$325,000

Redfin

### 3. MODELS

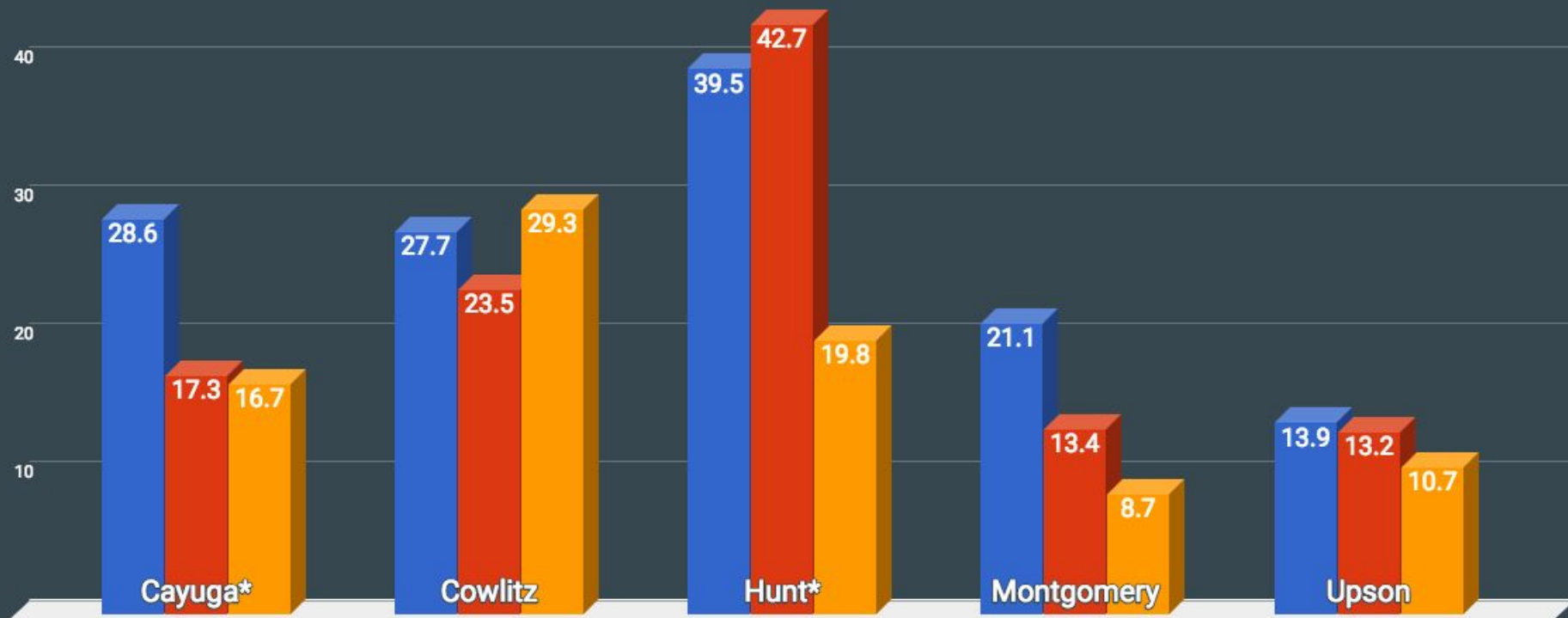
- **Linear Regression** (Baseline model):
  - Frequently used in Economics paper
- **Support Vector Regression** (SVR):
  - Good at finding signals and ignoring noises
- **Random Forest** (RF):
  - Good for datasets with missing values



# 3. RESULTS

## Percentage Of Houses In 5% Range

■ My Data's Baseline ■ SVR - Different Attributes ■ Zillow's Countywide Baseline





# 3. RESULTS

UPSON	COWLITZ	MONTGOMERY	HUNT	CAYUGA
Bed	Bed	Bed	Size	Size
Bath	Bath	Bath	Date Built	Tax Amount
Dishwasher	Asphalt Roof	Dishwasher	Tax Amount	Dishwasher
Assessment	Assessment	Assessment	Assessment	Assessment
Lot	Elementary-	Size	Hardwood-	Last
Date Built	School Score	Date Built	Floor	Remodel-Year
Walk Score	Size	Walk Score	Walk Score	Walk Score
...	...	...	...	...

# 3. RESULTS

## SAME ATTRIBUTE:

1. Bed
2. Bath
3. Dishwasher
4. Size
5. Tax Amount
6. Walk Score
7. Price Listed
8. Date Built
9. Assessment
10. Comparables' Sold Price

# 3. RESULTS

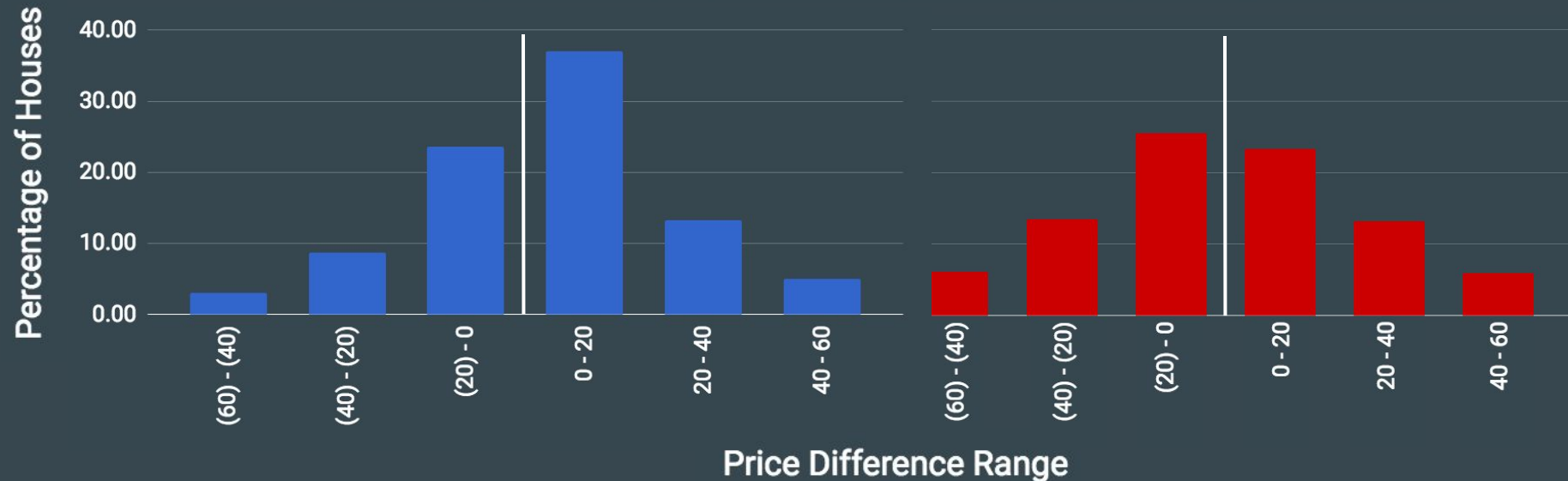
## Different Attributes vs Same Attributes

■ My Data's Baseline ■ SVR - Different Attributes ■ SVR - Same Attributes



# 3. RESULTS

## House Distribution In Price Difference Range



Overestimated : Underestimated Ratio = 3:2  
Zillow

Overestimated : Underestimated Ratio = 1:1  
My Predictor

### 3. RESULTS

- Weights of Important Attributes Across 5 Counties:
  - \$1 increase in Tax Assessment increases Sold Price by 54 cents
  - \$1 increase in Comparables' Sold Price increases Sold Price by 34 cents
  - \$1 increase in Price Listed increases Sold Price by 38 cents
  - 1 more Bathroom increases Sold Price by \$15,787

## 4. CONCLUSION

- Can I get close to or beat the Zestimate?
  - Beat Hunt's accuracy score and come close to Cowlitz's and Upson's.
- Can my models get rid of the overestimation problem?
  - Reduce the overestimated to underestimated ratio from 3:2 to 1:1
- Most important attributes?
  - Tax Assessment, Comparables' Sold Price, Price Listed, and Num of Bathrooms.

**THANK YOU FOR LISTENING!**