

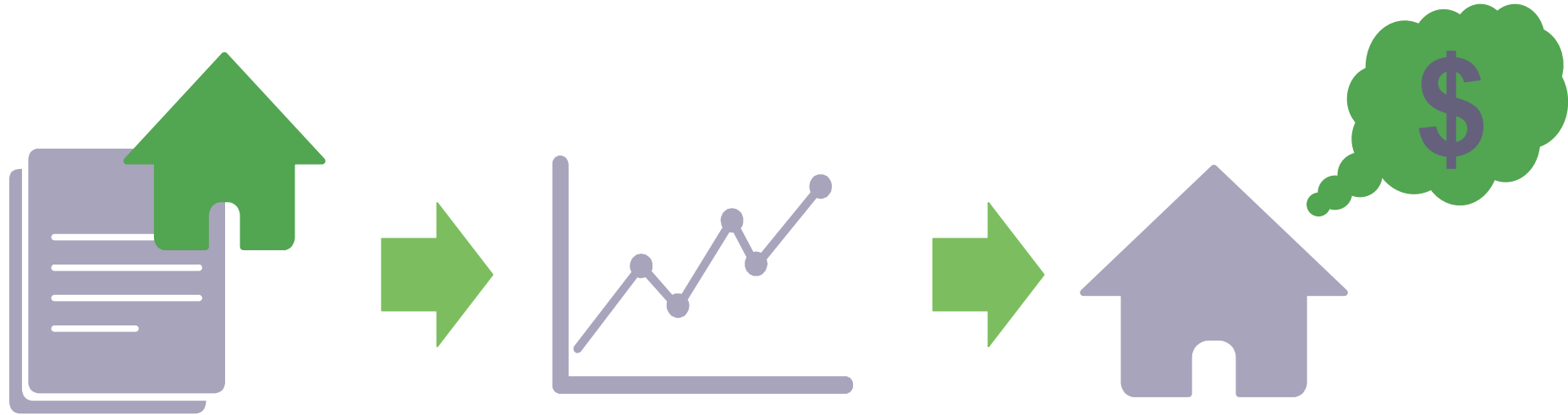
# Predicting House Sales



Natasha Kacoroski and David Kaspar

# How accurately can we predict the sale price of a house?

In one week, use housing data to build as accurate a predictive model as possible



# House Sales for King County

- Public domain
- **21,595** entries
- **2** years  
(May 2014-2015)
- **19** house features  
and price
- Engineered Features
  - Average zipcode price
  - Distance from Seattle



# House Sale Price Regression Model

Log of Predicted Values

14.0

13.5

13.0

12.5

12.0

11.5

12.0

12.5

13.0

13.5

14.0

Log of True Values

12.3% Mean  
Absolute Error

# Predictive Accuracy for House Sale Price

Predicted	Actual	Error
\$296,092	\$264,999	\$31,093
\$174,022	\$151,000	\$23,022
\$360,742	\$439,999	-\$79,257
\$336,684	\$272,925	\$63,759



# Moving Forward

## Model Refinement

1

Engineer nearest neighbor price feature

2

Apply model to full King County dataset

3

Further tune model parameters

