



# Model driven investing

---

IN KING COUNTY, WASHINGTON

BRIAN TRACY

DECEMBER 10, 2021

# Outline

---

- Business Problem
- Data & Methods
- Modeling Process
- Regression Results
- Recommendations
- Next Steps
- Questions

# Business Problem

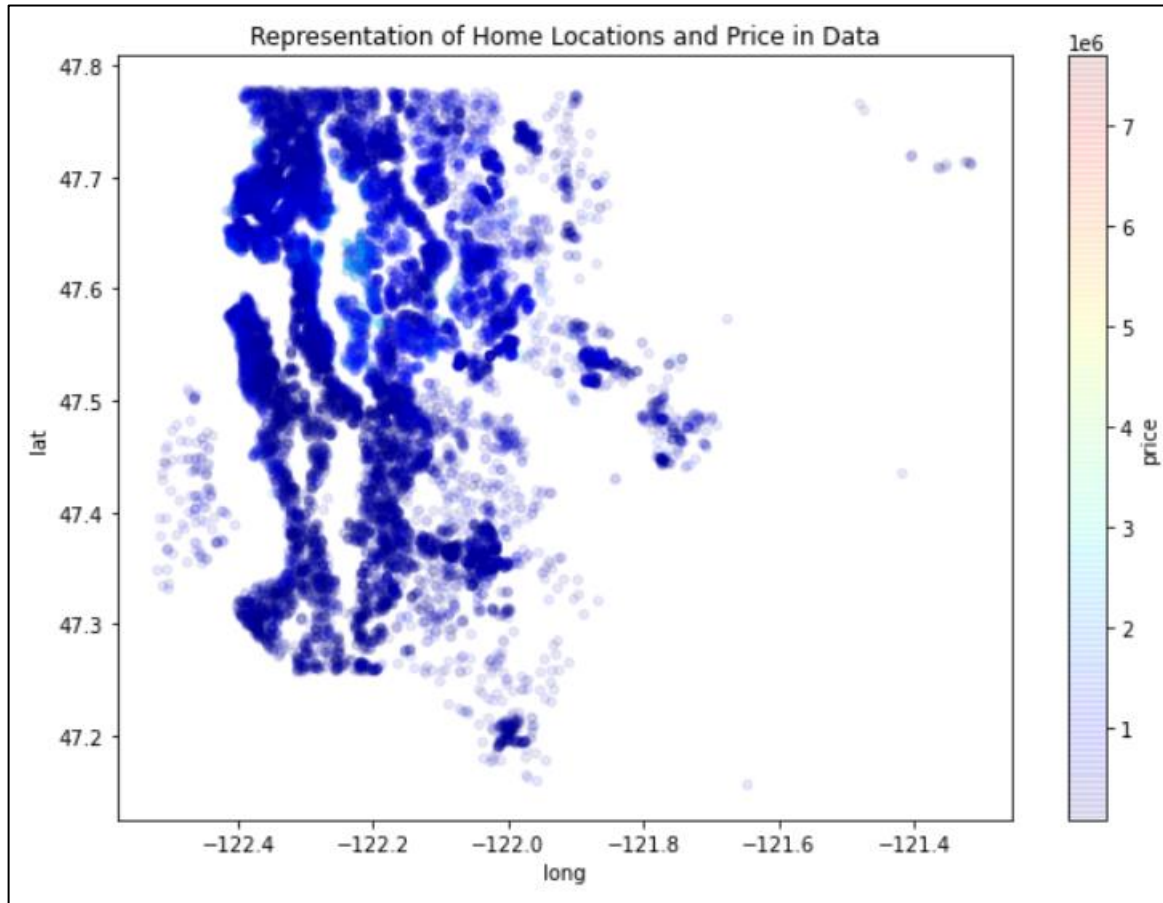
---

- A new bitcoin millionaire in Seattle has decided they want to diversify into real estate
- They want our company to build a model to predict home sales in King County, Washington
- The goal is to use our model to predict the price of a home currently for sale in order to influence the decision to buy.



# Data & Methods

---



- Data provided on home sale information from homes sold in King County between May 2014 and May 2015
- 21,597 entries in the data
- Features include:
  - Sale price
  - Home composition, condition and quality
  - Location
- Used python and Jupyter notebook for work environment

# Modeling Process

---

Model type:

- Linear Regression

Baseline Model:

- Used all available data

Successive iterations:

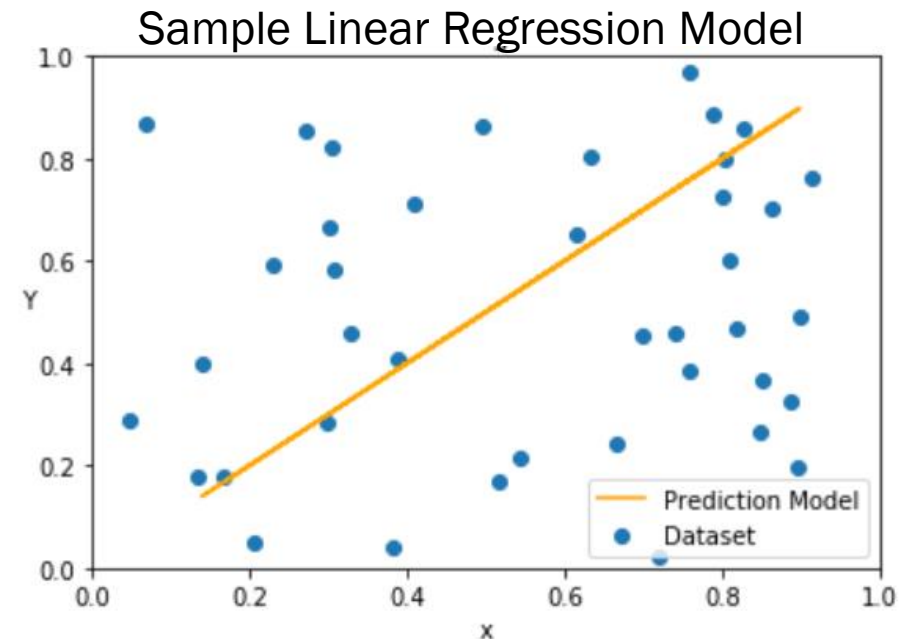
- Feature selection, engineering
- Model evaluation after each iteration

Testing the model:

- Train on 80% of the data, test on 20%

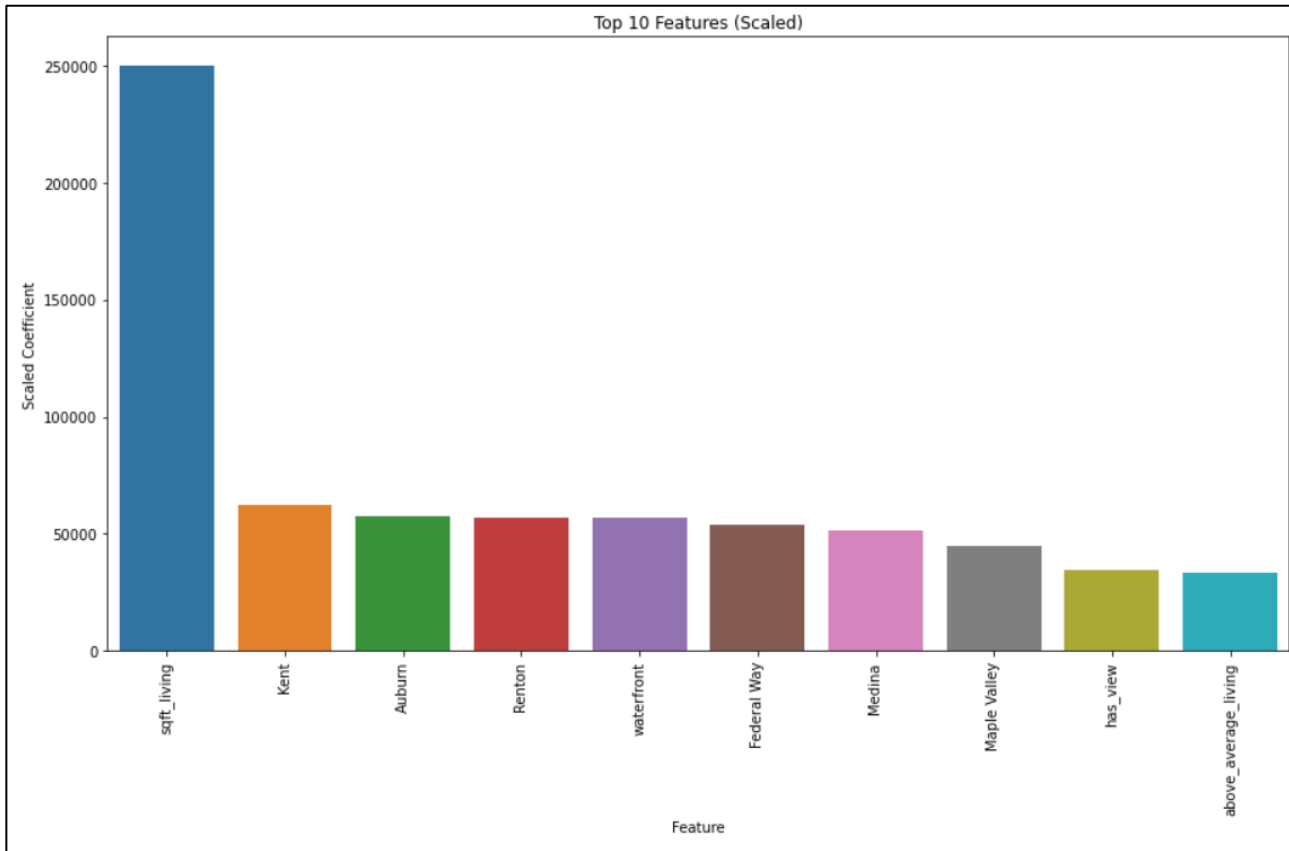
Scaling:

- In order to compare features



# Regression Results

---



- The final model can explain about 69% of the impact on a homes price based on the 32 different features used
- Square footage is the most influential feature on a homes price by a factor of 4
- Waterfront properties are on average 3 times more valuable than non-waterfront, and homes with a view adds about \$116k to value
- Location matters (relative to Seattle)
  - Home prices substantially drop in the towns just south of Seattle
  - The uber wealthy live in Medina

# Recommendations

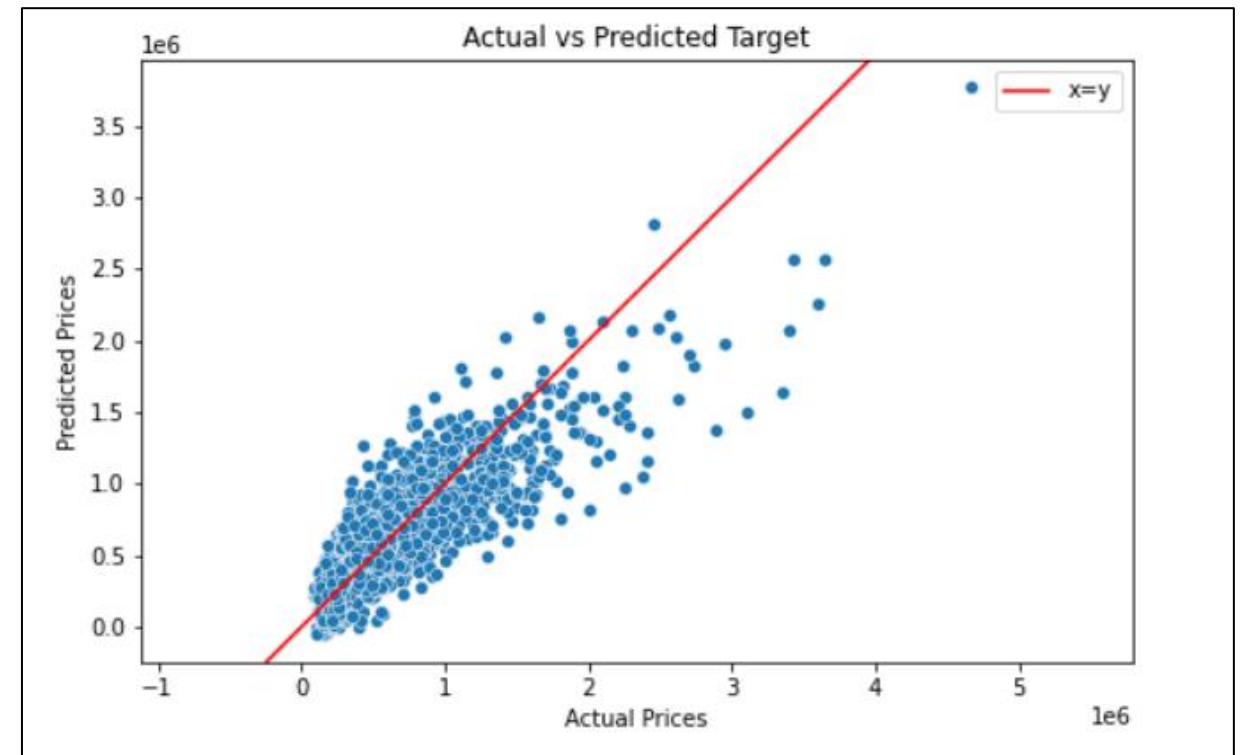
---

High error, use caution

- The model can predict home price with an error of \$127,175 USD; average home price in data is \$540,297 USD (error would be 23%)

Use model to advise what features to focus

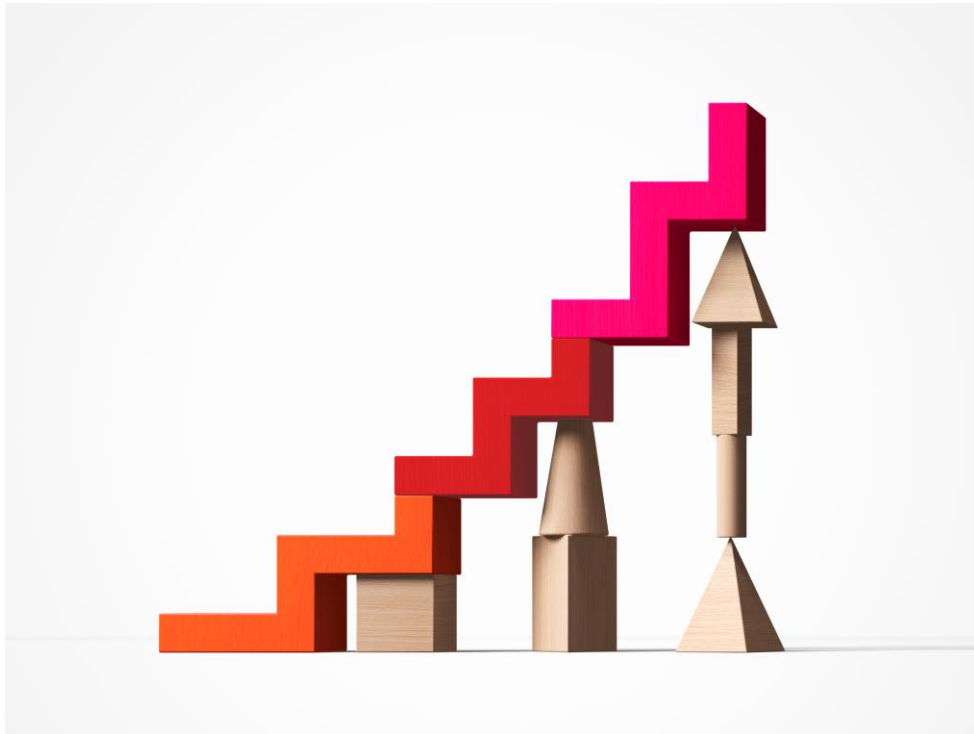
- Federal Way is the only city with lower-than-average home prices that also has waterfront property





# Next steps

---



So much a homes value is influenced by location. Use geospatial information to engineer more features:

- School districts, crime statistics, traffic/travel times, distance to shops are all factors

Most homes purchased as investments are rented. Get data on mortgage rates and rental prices to predict how much income a home could potential generate



# Questions?

---

GITHUB: [HTTPS://GITHUB.COM/BRTRACY/DSC-PHASE-2-PROJECT-V2-3](https://github.com/BRTRACY/DSC-PHASE-2-PROJECT-V2-3)