



# **Creating Accurate Real Estate Listings: The Power of Data**

**Peter Burton, Jillian Clark, Mychal Dubyk**



# Welcome King County Realtors!

**Business Problem:** Each new listing requires you to search through recent sales for “comparables”

Downsides to using comparables:

- Time consuming
- Inaccurate\*
- Based on past data\*

† <https://www.realtor.com/advice/buy/comps-are-off-base>

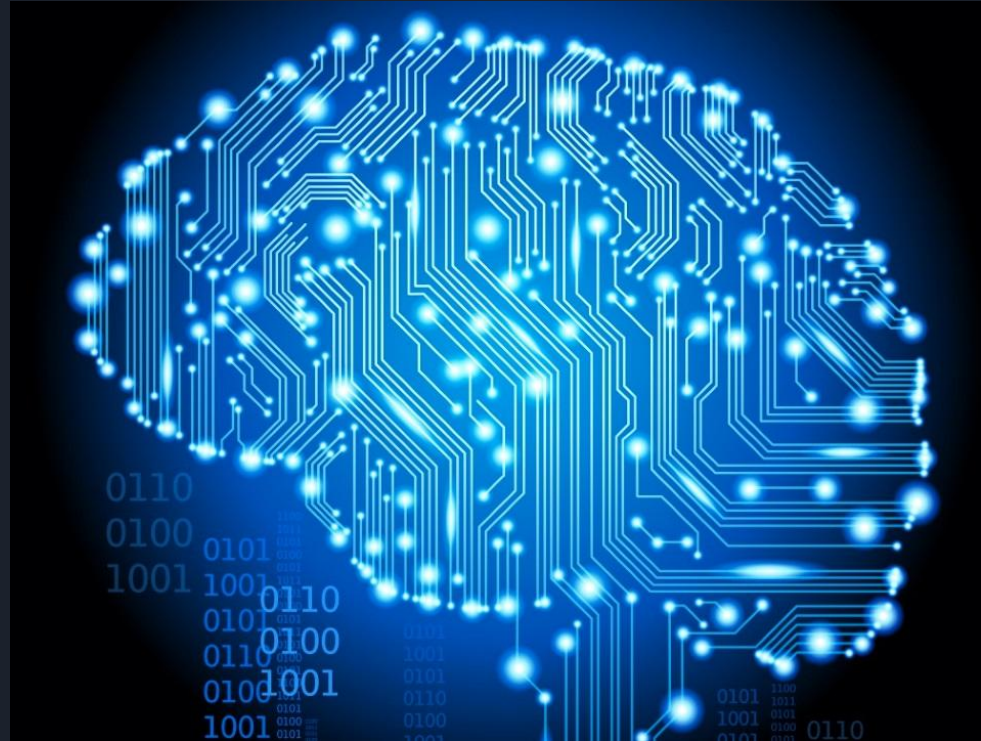
† Why you shouldn't base asking price on comparable sales - Inman

† <https://www.mansionglobal.com/articles/pricing-a-property-based-on-comps-is-still-the-rule-unless-it-s-incomparable-68596>

# Solution: Machine Learning Prediction Tools

## The solution?

*A data driven machine learning tool that can sift through large amounts of housing data and find trends, identify key features, and accurately and reliably predict the best listing price for a new listing in the King County housing market*





# Data Review and Key Variables In Our Model

Reviewed 21,597 listings from King County from May 2014-May 2015

Identified 3 key variables in the market data:

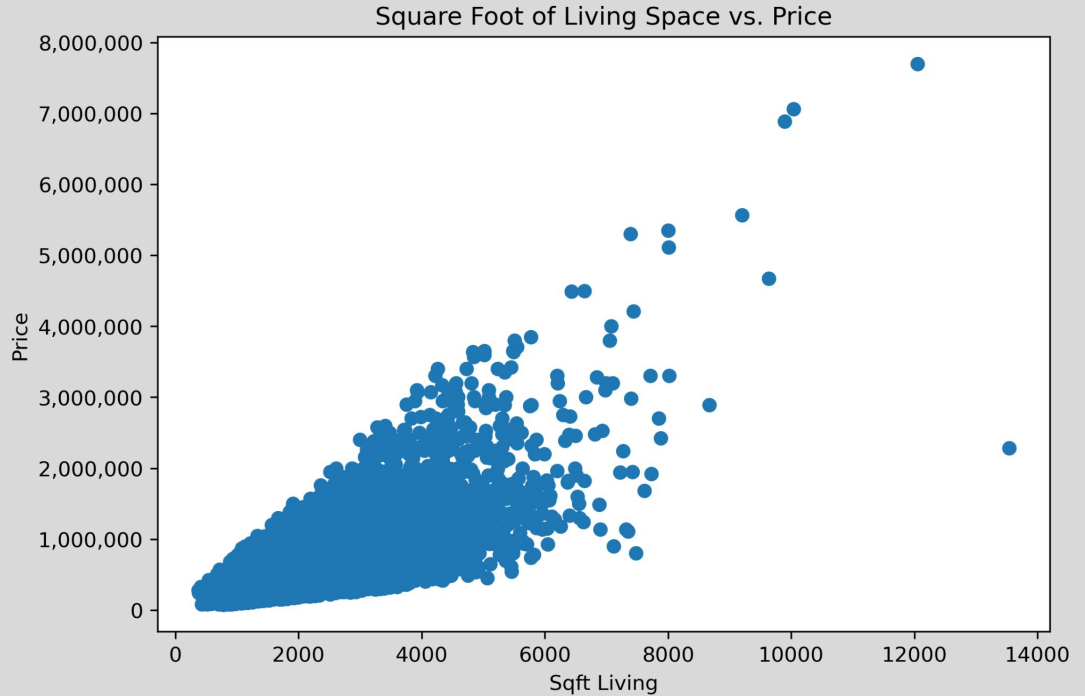
- Square Footage of Living Space
- Condition
- Waterfront Amenity

# Square Footage of Living Space

Living area increase of  
1000 square feet

=

Price increase of \$96,750

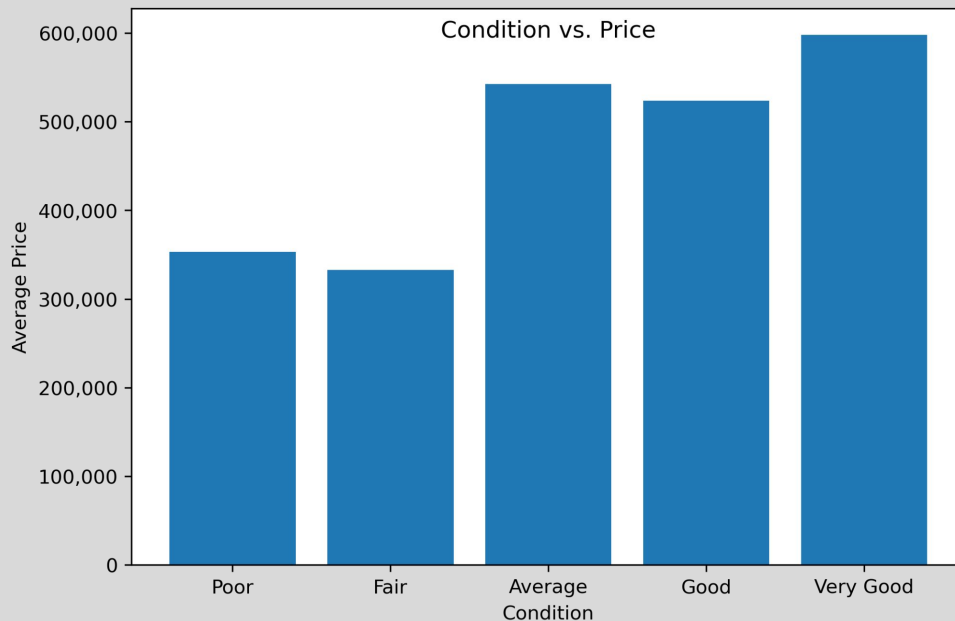


# Condition

-There is an average decrease of \$93.1K in Poor homes compared to Average homes

-There is an average increase of \$34.7K in Good homes compared to Average homes

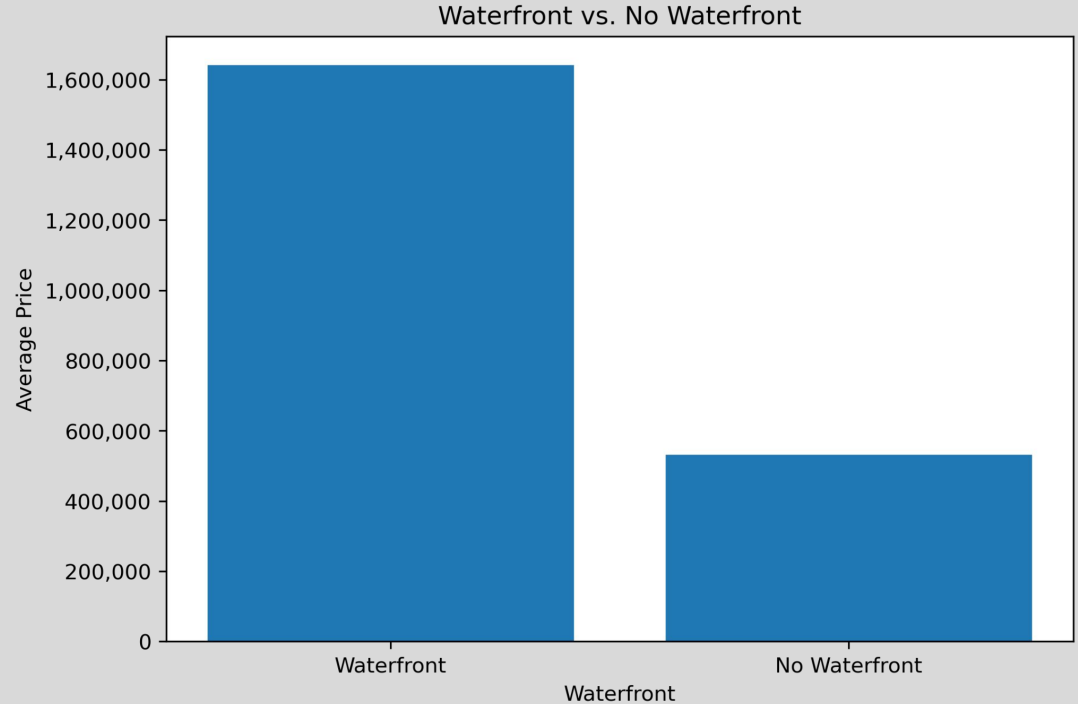
-There is an average increase of \$65.6K in Very Good homes compared to Average homes



# Waterfront Amenity

-Waterfront properties sell for an average of around 3x more than non-Waterfront

-Waterfront properties sell for \$510,400 more than non-Waterfront



# Conclusion

**Creating listing prices based on our model is a significant improvement**

**The model is more accurate then using comparables:**

- The model has a linear regression  $r^2$  of .74, and a polynomial  $r^2$  of .86
- Significant improvement compared to choosing 2-3 comparables or simple model (.49  $r^2$ )

**The predictions can be made quickly, and will save time**







## Next Steps

- Identify housing market trends over time, and bake that into listing prices
- Identify more key features and create a larger data set to create a more accurate model
- Create a similar tool for buyers

# Thank you! Any questions?

## The ABCD TEAM:

### Analysis by **B**urton, **C**lark, **D**ubyk

Peter Burton

E-mail:  
[akburton21@gmail.com](mailto:akburton21@gmail.com)

[GitHub](#)

Jillian Clark

E-mail: [jlc0512@gmail.com](mailto:jlc0512@gmail.com)

[GitHub](#)

Mychal Dubyk

E-Mail:  
[mychaldubyk@gmail.com](mailto:mychaldubyk@gmail.com)

[GitHub](#)