

# King County Real Estate Report

*by Doug and the Data Diggers*



# Intros



**Doug Mill**

Data Scientist

Proprietary Trader &  
Investor, BA



**Chris O'Malley**

Data Scientist

Electronic Technician



**Jakub Rybicki**

Data Scientist

Geotechnical  
Engineer, MS





# Agenda

1

Business Understanding

2

Data Understanding and Preparation

3

Modeling

4

Evaluation

5

Proposal

6

Future Investigations

# Real Estate Agency



- Important Home Characteristics
- Using the 2015 Housing Market

Business

Data

Modeling

Evaluation

Proposal

Future



**More Square footage**

**Better school district**

**Higher Grade**

**=**

**MORE  
MONEY**



**Business**

**Data**

**Modeling**

**Evaluation**

**Proposal**

**Future**





Shaded areas indicate U.S. recessions.

Source: S&P Dow Jones Indices LLC

fred.stlouisfed.org

S&P Dow Jones Indices LLC, S&P/Case-Shiller WA-Seattle Home Price Index [SEXRNSA], retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/SEXRNSA>, October 6, 2021.

Business

Data

Modeling

Evaluation

Proposal

Future



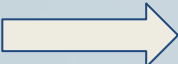


# Investor-driven Data

- Home sales in King County, Washington
- Sales between May 2014 - May 2015
- Data from 21.6K home sales
- 2015 School GIS data from [kingcounty.gov](http://kingcounty.gov)



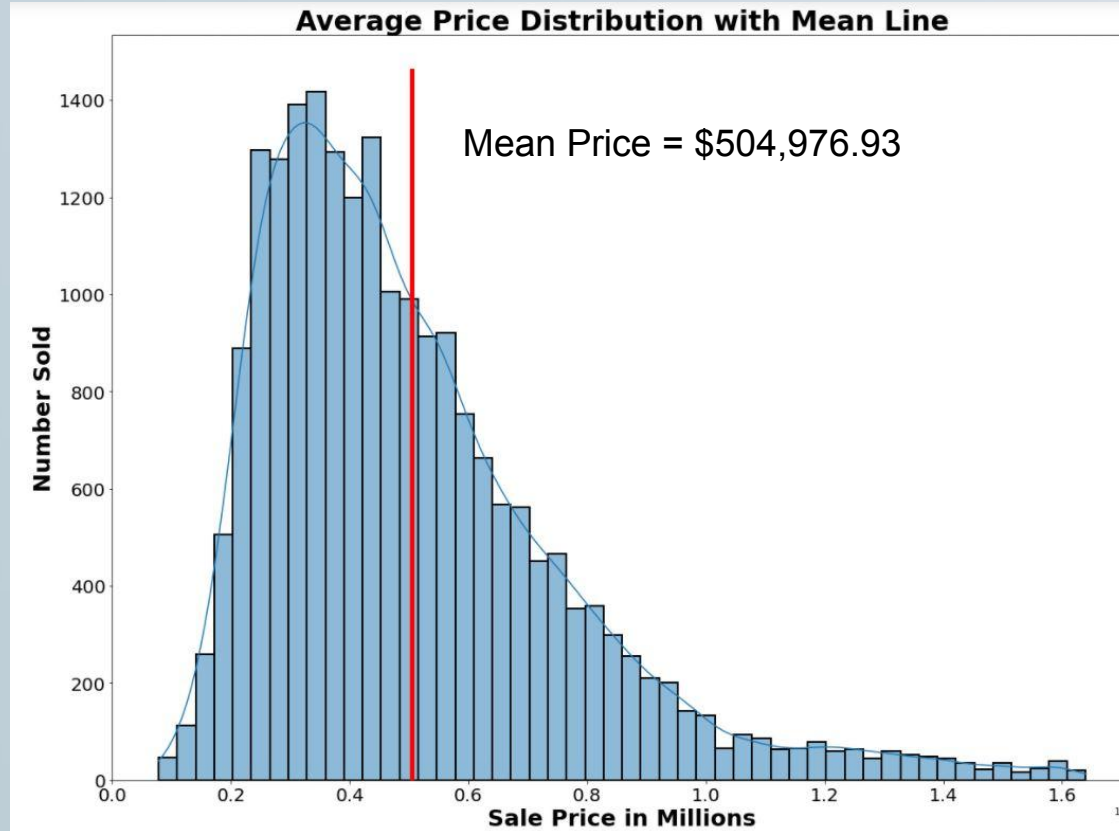
# Data Preparation

- Eliminated homes above \$1.65M
- Removed ~400 outliers and typos
- Data from 21.6k  21.2k home sales





# Exploring Data



Business

Data

Modeling

Evaluation

Proposal

Future



# Why so Pricey?

?

?

?

What are the key variables driving home values?

Business

Data

Modeling

Evaluation

Proposal

Future



# Key Factors

1



Home Size

2



School District

3



Grade

Business

Data

Modeling

Evaluation

Proposal

Future



# Why School District?

School quality and district can greatly impact home value



Business

Data

Modeling

Evaluation

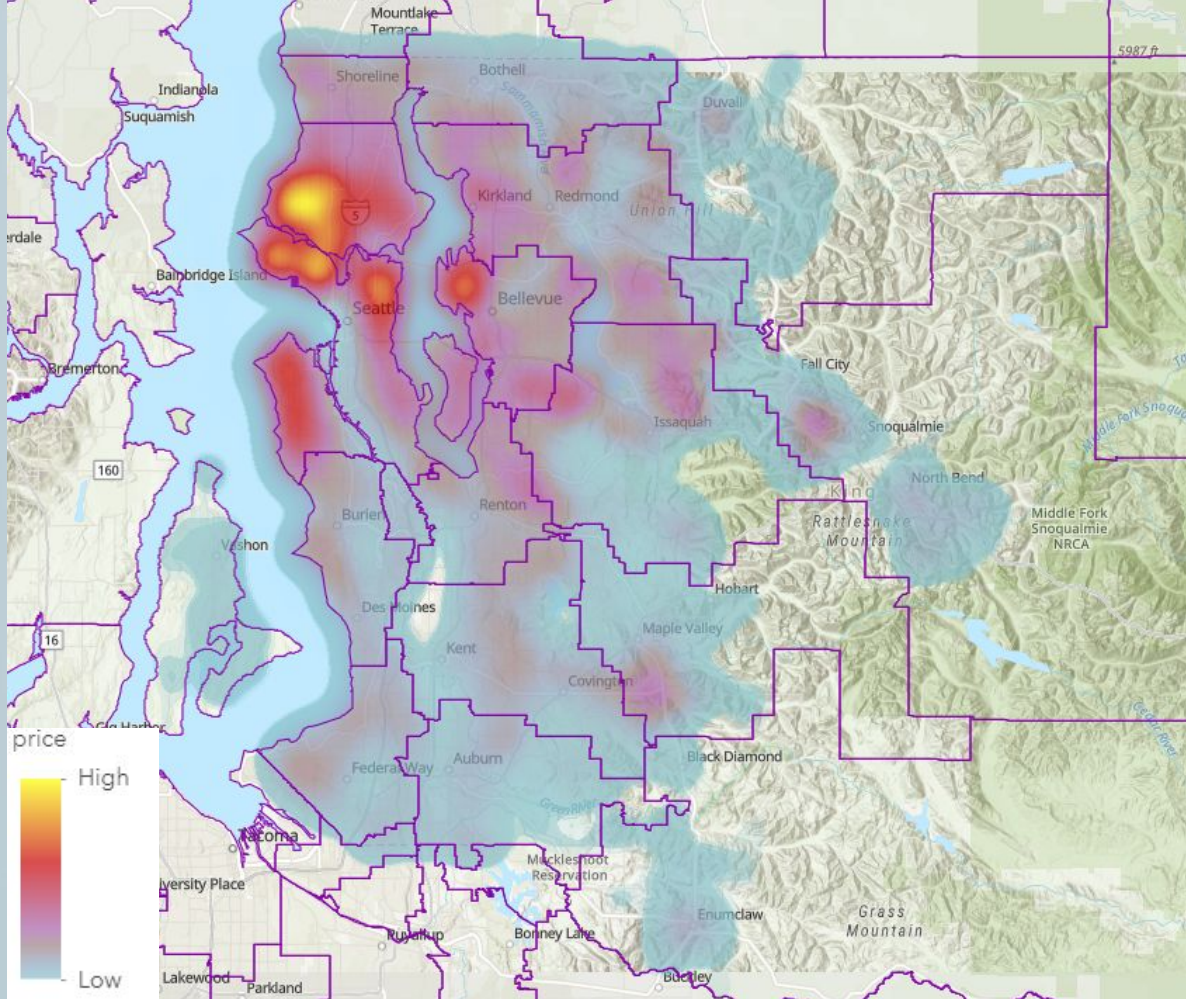
Proposal

Future



## TOP 6 School Districts

School District	Mean Price
MERCER ISLAND	\$1,002,759.00
BELLEVUE	\$747,735.80
ISSAQUAH	\$675,401.00
LAKE WASHINGTON	\$631,370.30
SEATTLE	\$557,237.90
NORTHSHORE	\$531,400.90



Business

Data

Modeling

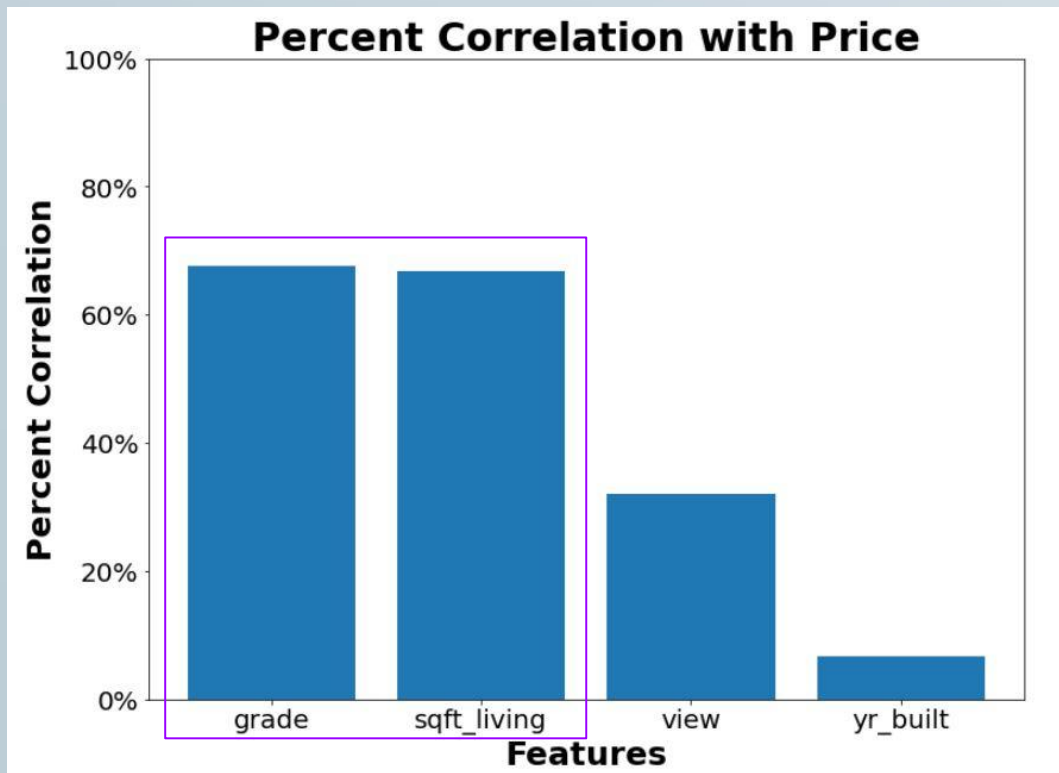
Evaluation

Proposal

Future



# Modeling



Business

Data

Modeling

Evaluation

Proposal

Future



# Evaluation

<u>Variable</u>		<u>Effect on Home Price</u>
Home Size		\$235 /sqft
School Districts (Top 3):	Mercer Island District (1)	+\$508,808
	Bellevue District (2)	+\$360,100
	Seattle District (3)	+\$297,400
Building Grade		\$71,090 /grade

Business

Data

Modeling

Evaluation

Proposal

Future





# Proposal



- Low square footage
- Federal Way, Kent, Tukwila
- Grade < 7



- High square footage
- Mercer Island, Bellevue, Seattle
- Grade > 9

Business

Data

Modeling

Evaluation

Proposal

Future



# Future Investigations

- Proximity to shopping, entertainment, amenities, employment, and airports
- Compare sales by neighborhood



Thank  
You



**Doug Mill**

[github.com/cyborgsage](https://github.com/cyborgsage)



**Chris O'Malley**

[github.com/chriso2672](https://github.com/chriso2672)



**Jakub Rybicki**

[github.com/Jakubryb](https://github.com/Jakubryb)



Questions?



# Appendix



School District	Mean Price
MERCER ISLAND	\$1,002,759.00
BELLEVUE	\$747,735.80
ISSAQUAH	\$675,401.00
LAKE WASHINGTON	\$631,370.30
SEATTLE	\$557,237.90
NORTHSHORE	\$531,400.90
VASHON ISLAND	\$485,801.80
SNOQUALMIE VALLEY	\$484,465.40
RIVERVIEW	\$442,684.40
SHORELINE	\$423,701.10
TAHOMA	\$381,659.30
RENTON	\$362,342.50
KENT	\$320,962.40
ENUMCLAW	\$317,884.10
HIGHLINE	\$317,111.70
FEDERAL WAY	\$290,356.70
AUBURN	\$290,268.10
TUKWILA	\$249,271.10
SKYKOMISH	\$218,562.50

Business

Data

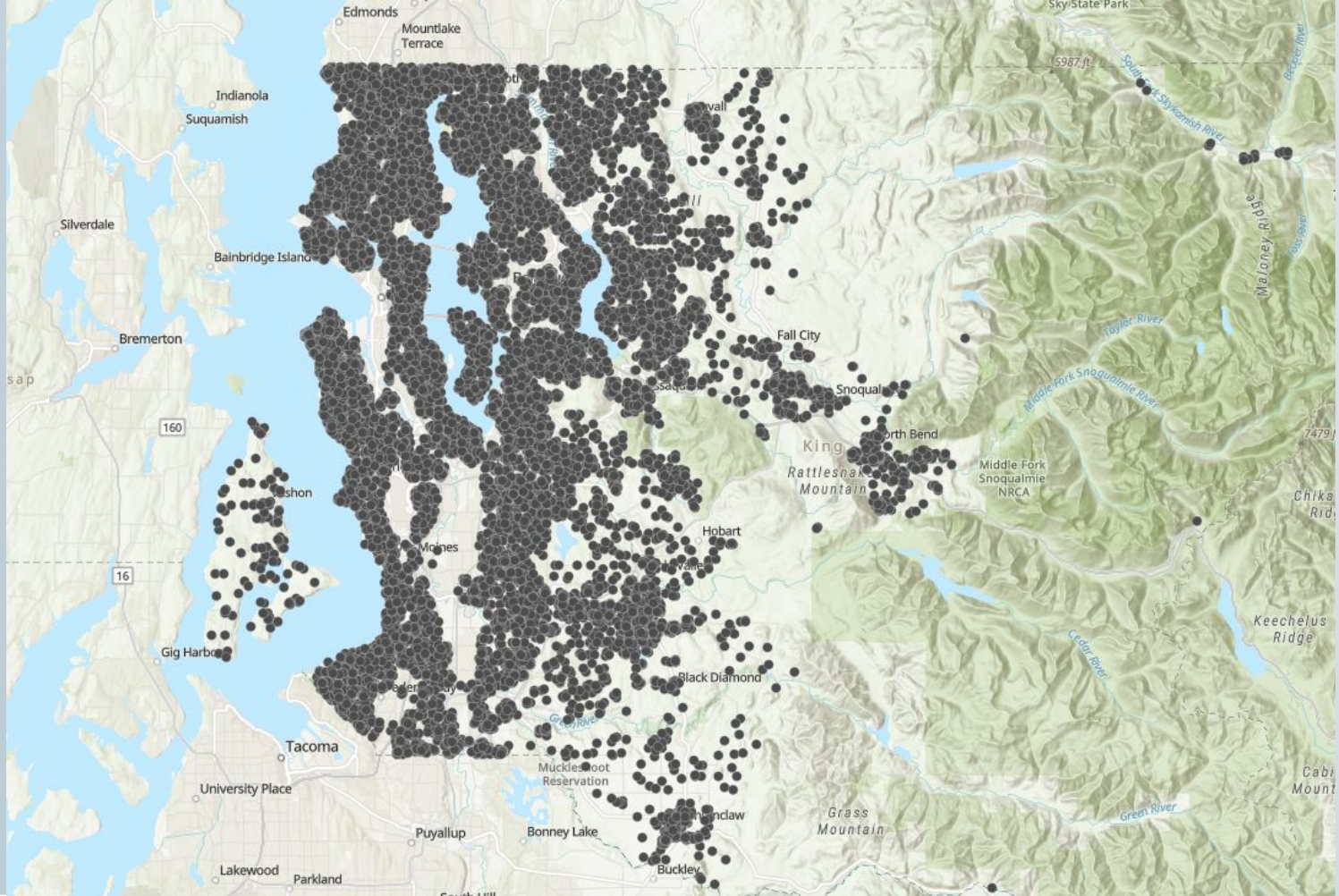
Modeling

Evaluation

Proposal

Future





Business

Data

Prep.

Modeling

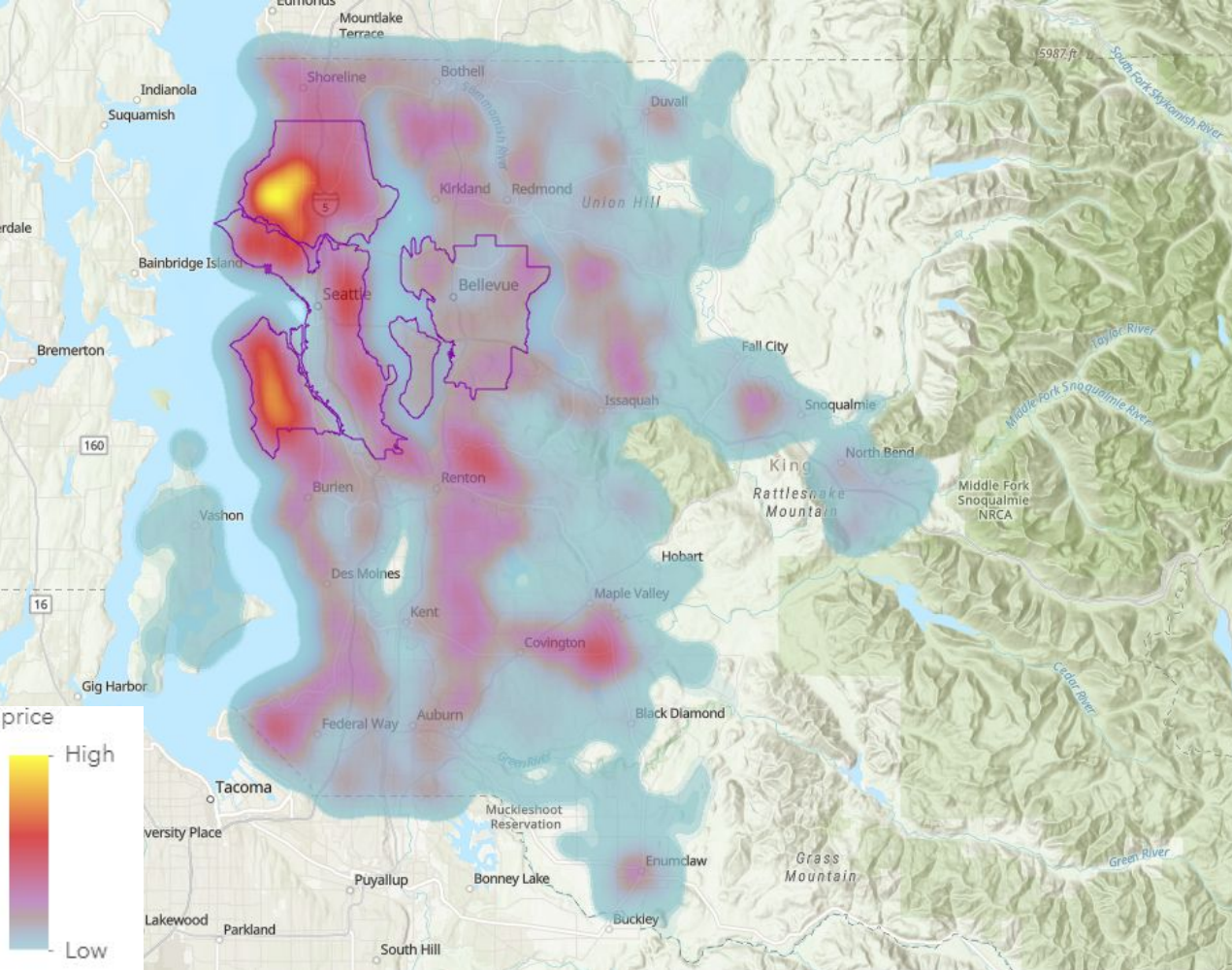
Evaluation

Proposal

Future







## Top 3 School Districts

Business

Data

Prep.

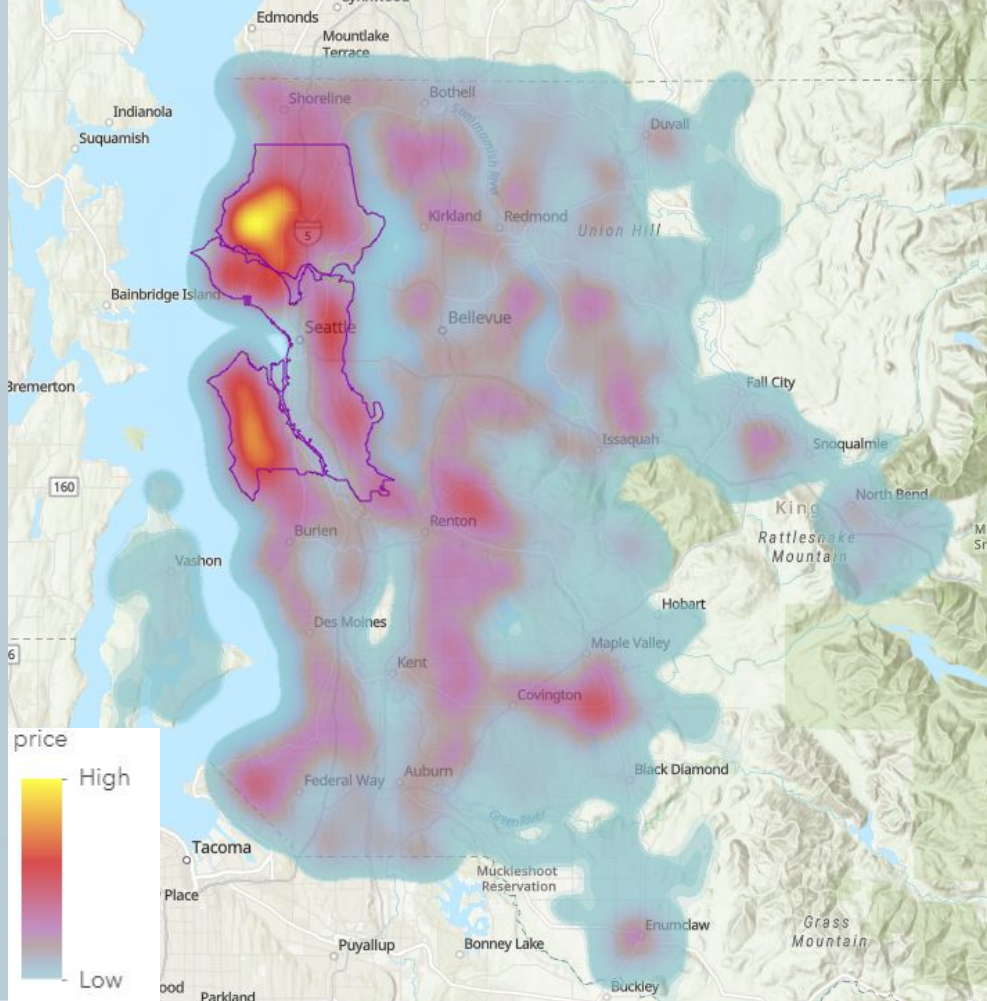
Modeling

Evaluation

Proposal

Future





## Public School District

Business

Data

Prep.

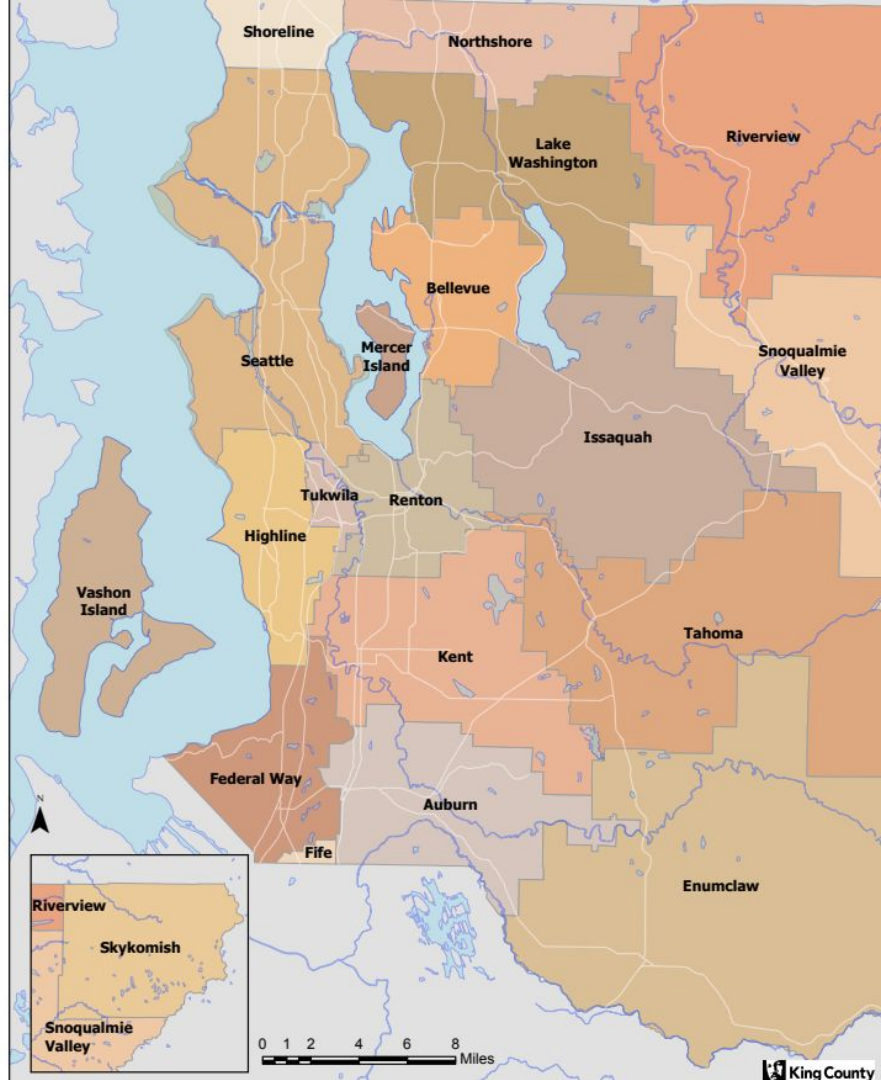
Modeling

Evaluation

Proposal

Future



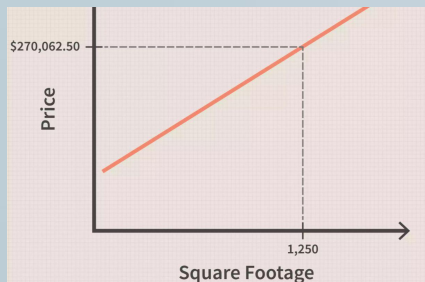


# King County School Districts



# Assumptions of Linear Regression

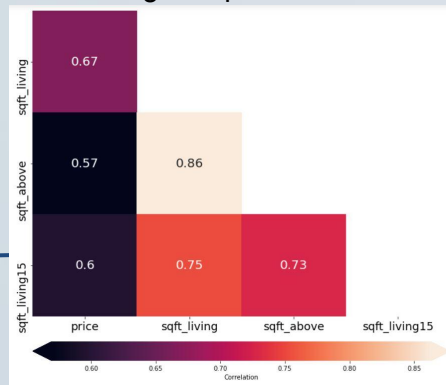
Straight-line Relationship



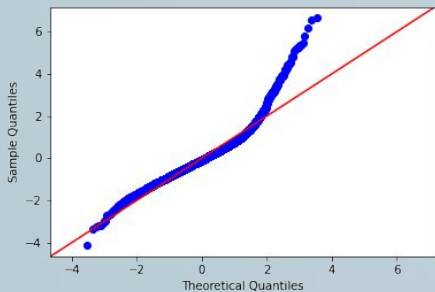
**Linear Relationship**

**Low Multicollinearity**

Validating Independence



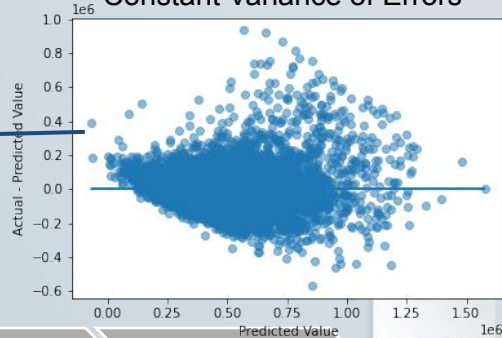
Error Distribution



**Normal Distribution of Errors**

**Homoscedasticity of Errors**

Constant Variance of Errors



Business

Data

Prep.

Modeling

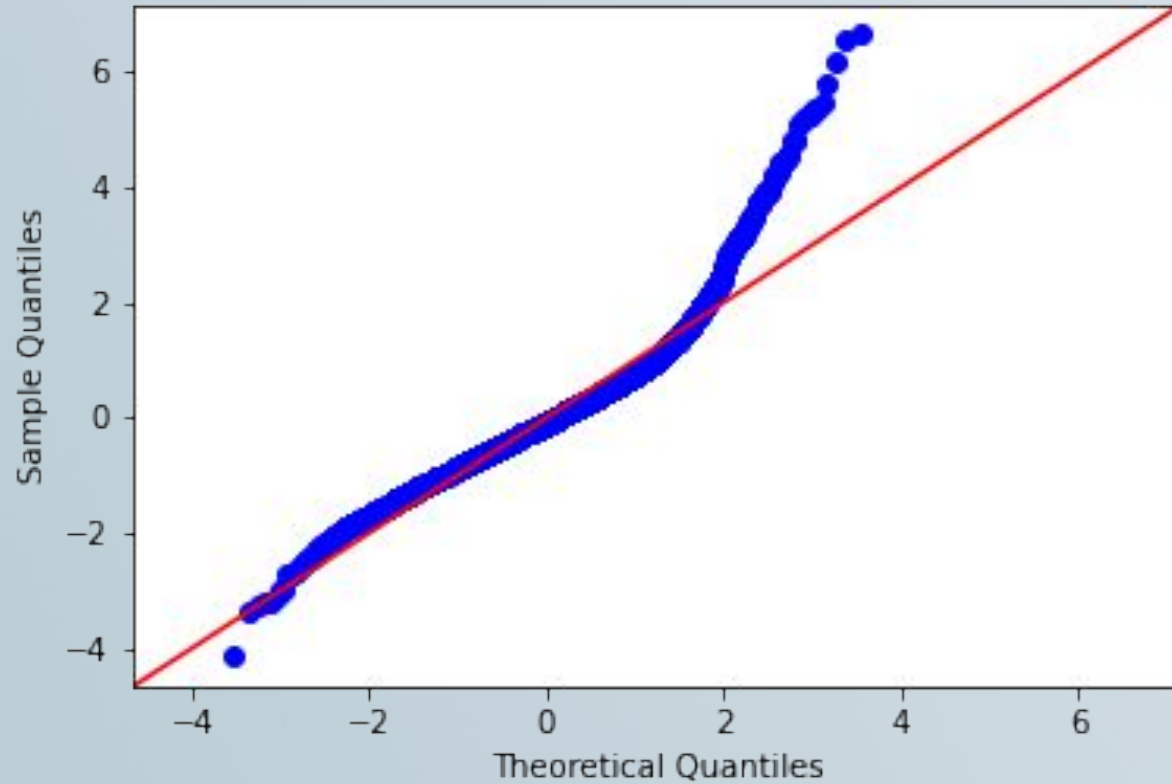
Evaluation

Proposal

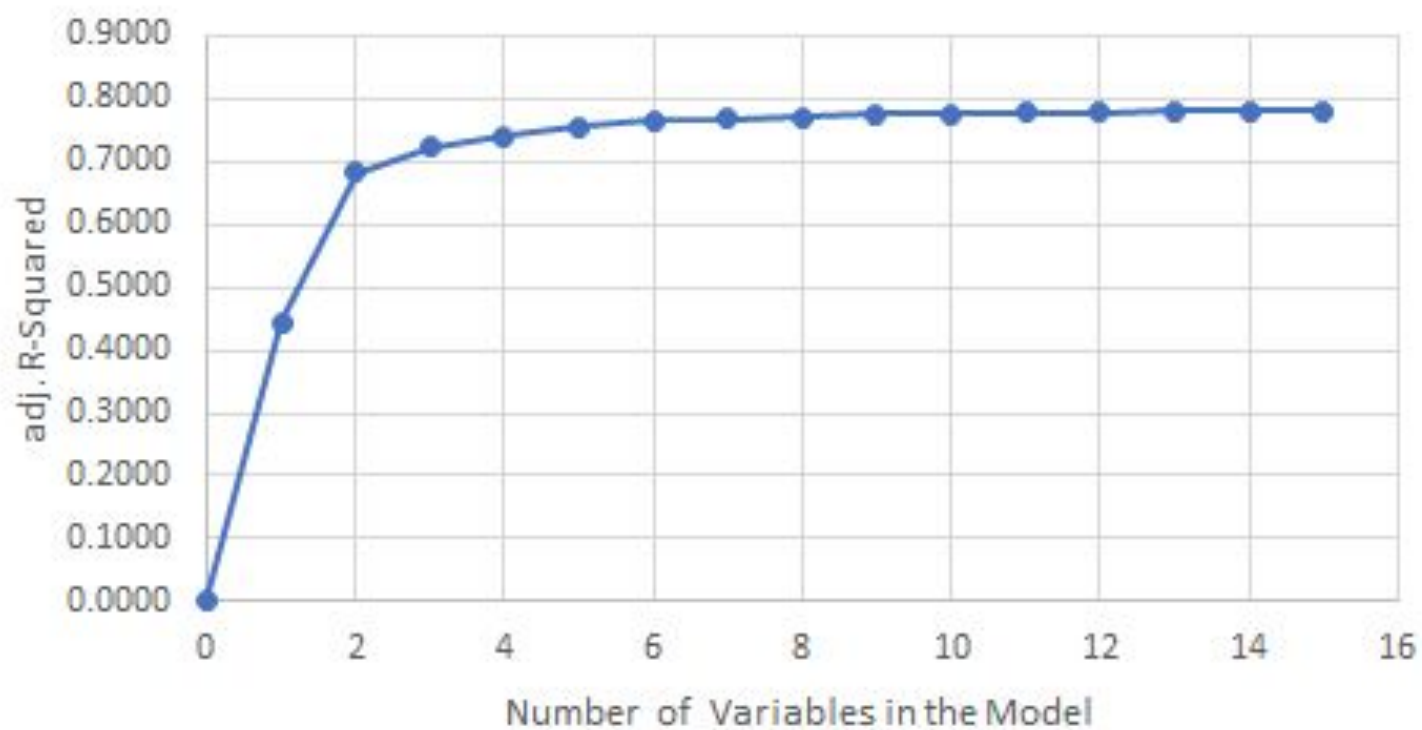
Future



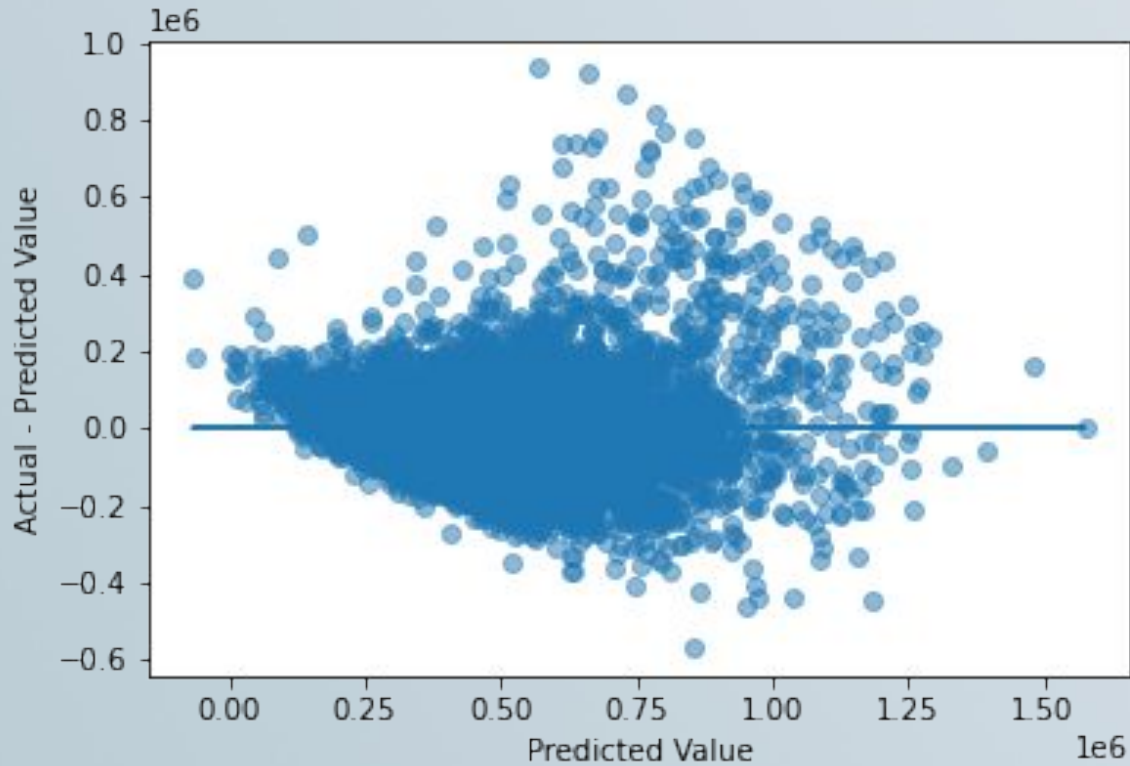
## Normal Distribution of Errors



Rval vs # of Variables



## Homoscedasticity of Errors





# This Chart Shows Why This Rally Has Been So Hated

Myles Udland Jun 17, 2014, 10:32 AM

An overriding market meme since the beginning of 2013 is we are in the midst of the most hated bull market in history.



<https://www.businessinsider.com/hated-bull-market-2013-2014-2014-6>  
<https://www.macrotrends.net/2604/30-year-fixed-mortgage-rate-chart>

Business

Data

Prep.

Modeling

Evaluation

Proposal

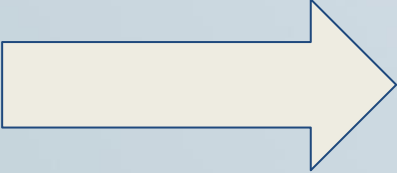
Future



# Improving our models

## Variable

## Explainability

1. Sqft_living		44.5%
2. School District		68.3%
3. Grade		72.3%
4. Year Built		74.2%

Business

Data

Prep.

Modeling

Evaluation

Proposal

Future



**PREPARING  
THE BLUFF**



*PLEASE WAIT...*

