



# MODULE 1 PROJECT

King County Housing  
By Desmond Webb

# King County Housing

- 21597 entries
- 21 variables
- May 2014 to May 2015
- Living space
- Lot space
- Neighborhood scale

# Pricing Model

- Purpose:
  - *Predicting home price within King county*
  - *Exploratory data analysis*
- Predictors
  - *14 chosen*
- Hypothesis: Home price is predicted by selected predictors

# Findings

- Hypothesis: supported
  - *p-value can be trusted with caution*
- Large conditional number
  - *too much noise*
- Positive trend (price)

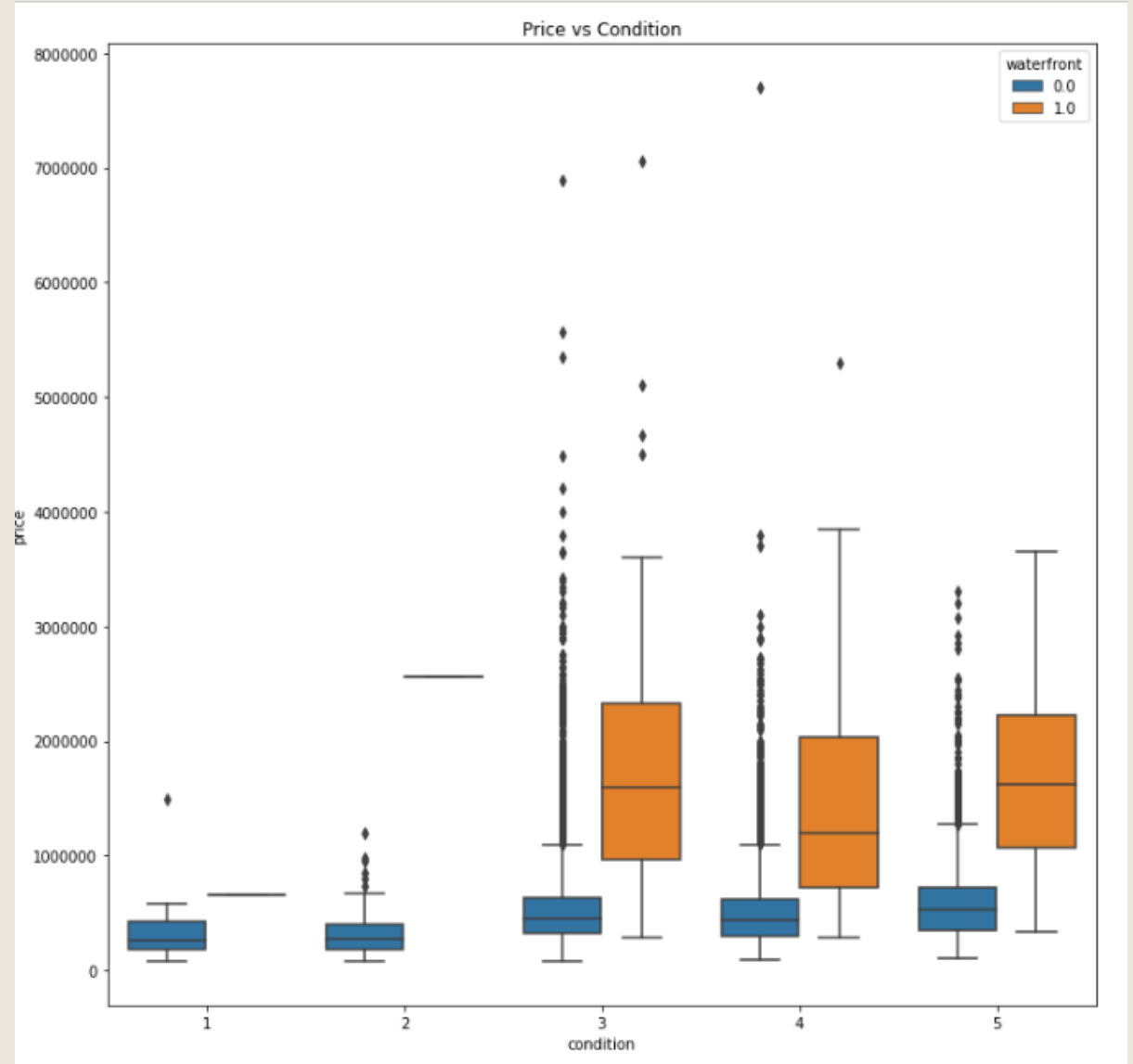
## OLS Regression Results

<b>Dep. Variable:</b>	price	<b>R-squared:</b>	0.665
<b>Model:</b>	OLS	<b>Adj. R-squared:</b>	0.665
<b>Method:</b>	Least Squares	<b>F-statistic:</b>	3058.
<b>Date:</b>	Sun, 07 Jul 2019	<b>Prob (F-statistic):</b>	0.00
<b>Time:</b>	18:23:36	<b>Log-Likelihood:</b>	-2.9559e+05
<b>No. Observations:</b>	21597	<b>AIC:</b>	5.912e+05
<b>Df Residuals:</b>	21582	<b>BIC:</b>	5.913e+05
<b>Df Model:</b>	14		
<b>Covariance Type:</b>	nonrobust		

<b>Omnibus:</b>	15149.875	<b>Durbin-Watson:</b>	1.995
<b>Prob(Omnibus):</b>	0.000	<b>Jarque-Bera (JB):</b>	809022.411
<b>Skew:</b>	2.790	<b>Prob(JB):</b>	0.00
<b>Kurtosis:</b>	32.460	<b>Cond. No.</b>	2.40e+08

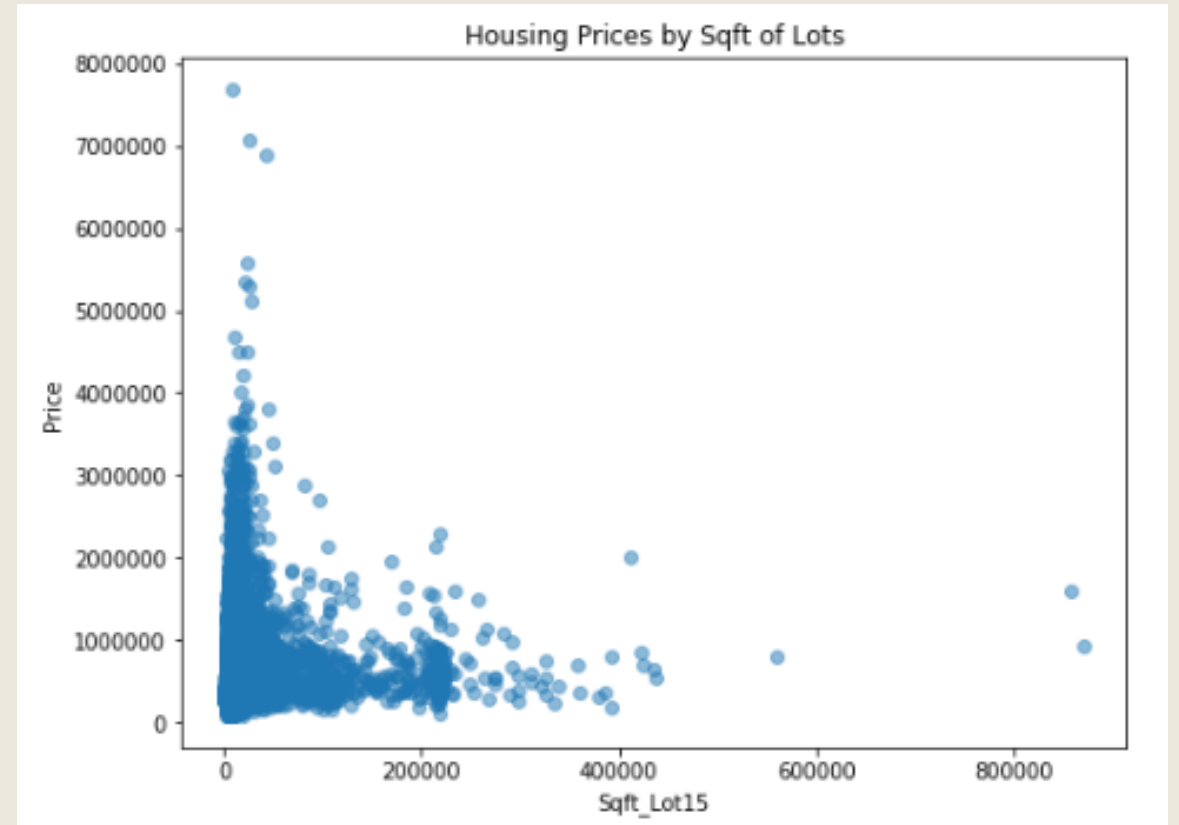
# Findings Cont...

- Price of Condition 3/5 homes
  - *Median*
- Condition 3 homes variance



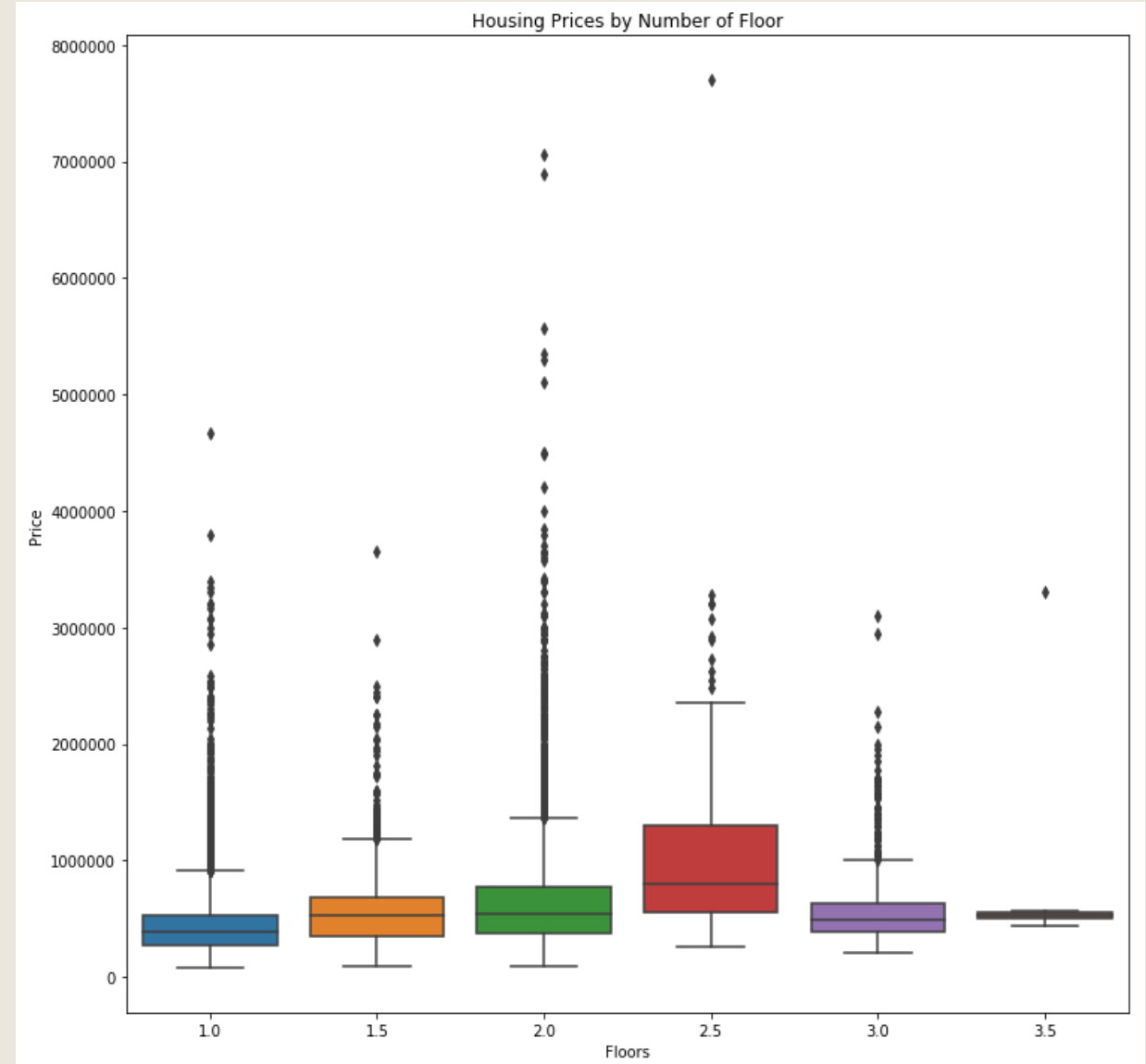
# Findings Cont...

- Cumulative lot size does not predict well



# Findings Cont...

- Number of floors predicts price...to a degree



# How We Improve

- Data acquisition
  - *Data custodian*
  - *Cross referencing*
    - view
    - Waterfront
- Model has potential iterability



# What Now?

- Develop a better model
  - *Additional multivariate techniques*
    - logistic regression analysis
    - MANOVA

Thank You