

Poverty Rates and Evictions in King County, WA

Team BCJ (#3)

Ben Kaliczak, Chathurangi Pathiravasan, Johann
Thiel

Goal(s)

- To study poverty rates and homelessness in King County, WA.
- To study evictions in King County, WA.
- Analyze and predict median housing value for King County, WA.

Hackathon

Modeling

REQUIREMENTS

Test assumptions

Test limitations

Models

Random Forest

GLM

GAM

Questions

What are the contributing factors which might have an effect for predicting median housing price

Is there a significant difference
between city center houses and
non city center houses

Can we accurately predict the
median housing price over the
next few years?

(Time Series)

Exploratory Data Analysis needed

Other variables

Demographics

Normalized ZHI - ZRI

Compared against evictions

Visualizations

Packages

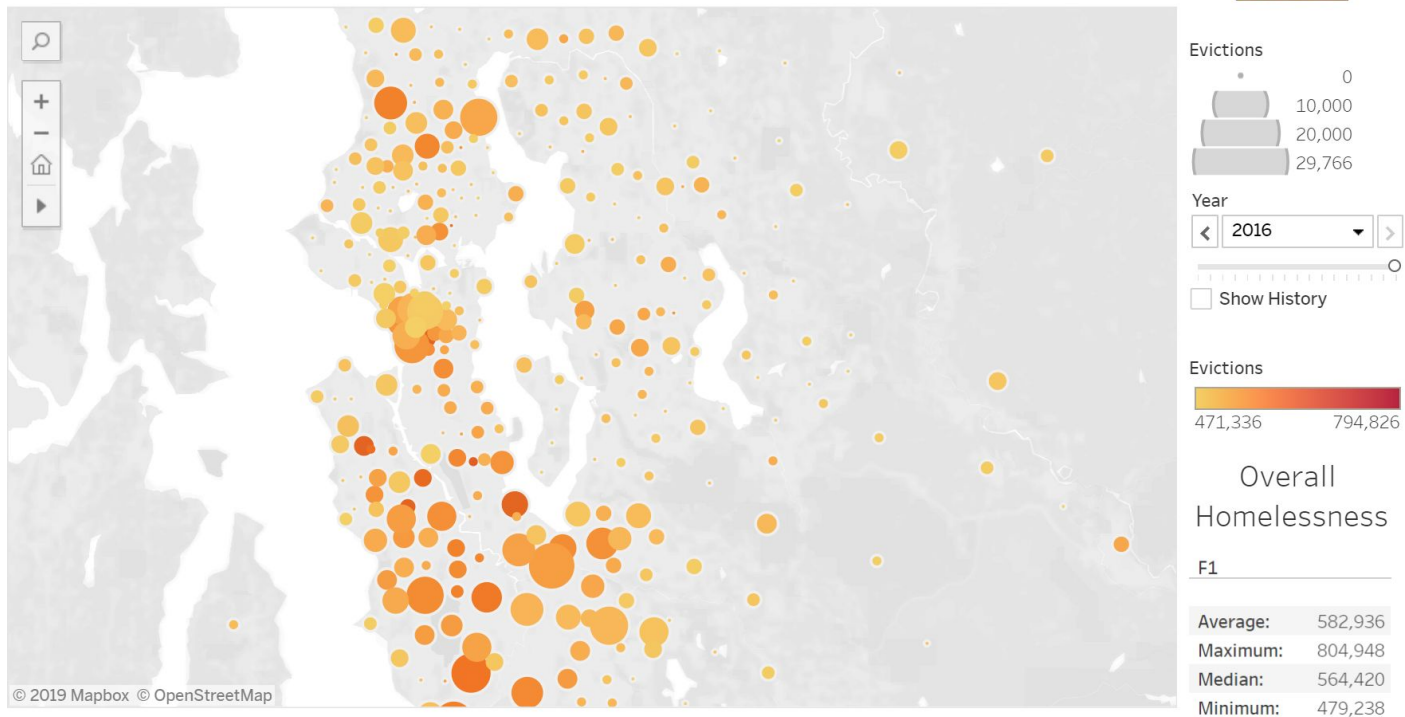
Mapview

Types of Visuals

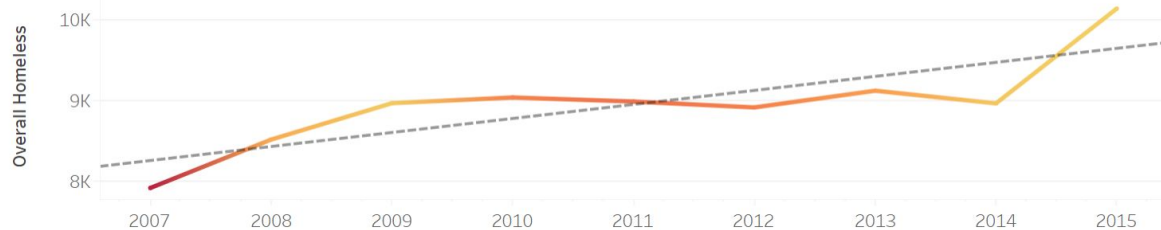
Homelessness and Poverty Rates

- Are homelessness counts higher in years with higher poverty and eviction rates?
 - *Through 2016 there was a 35% increase in exits from homeless shelters*
 - *Given the shortage of affordable housing options, the performance increase of the crisis response system is unlikely to sustain—there are fewer and fewer units available to house people. - McKinsey & Co*
- Data sourced from Hud Exchange and Eviction Lab
 - Evictions by GEOID within King County
 - Total evictions across King County
 - Poverty rates by GEOID within King County
 - Overall Homelessness counts across Seattle/King County
 - Overall Unsheltered Homeless counts across Seattle/King County
- [Poverty and Homelessness Tableau](#)

Poverty Rate by Evictions - 2016



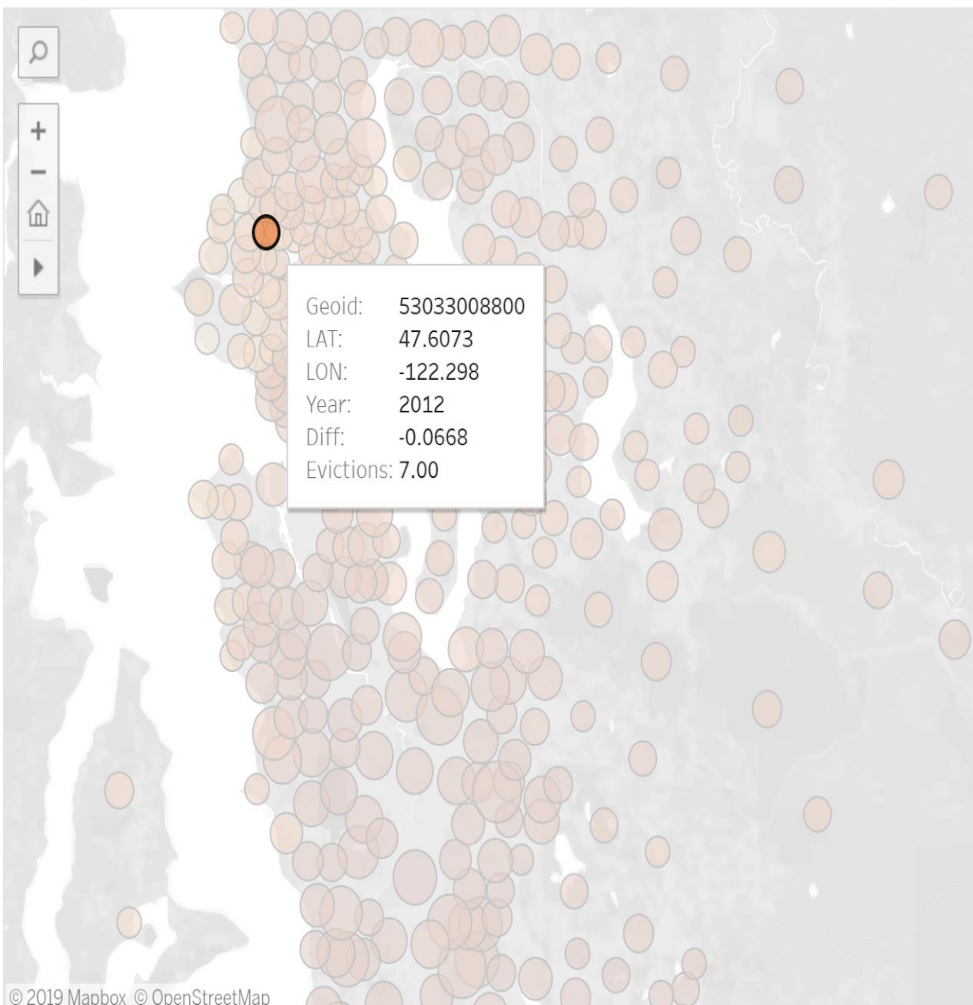
Overall Homelessness in King County



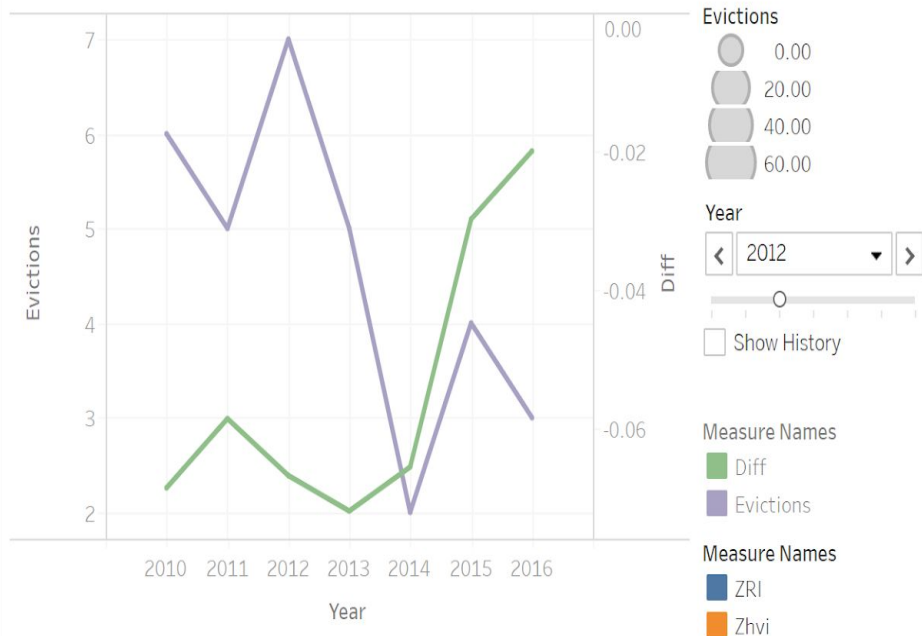
Evictions

- Are evictions more common in areas where the Zillow Home Value Index (ZHVI) is high or where the Zillow Rent Index (ZRI) is high?
- Data sourced from Zillow and Eviction Lab
 - Evictions by GEOID across King County
 - Crude measure: $\text{Diff} = \text{ZHRI (normalized by year)} - \text{ZRI (normalized by year)}$.
 - Diff by GEOID across King County
- [Evictions Tableau](#)

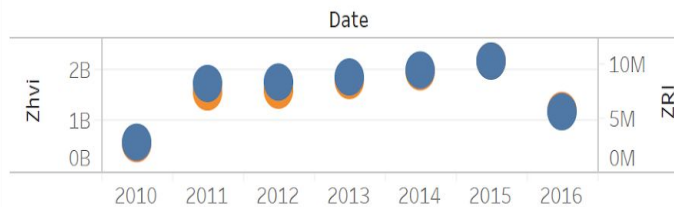
Evictions and Diff by GEOID for a fixed year



Evictions and Diff from 2010-2016 for a fixed GEOID



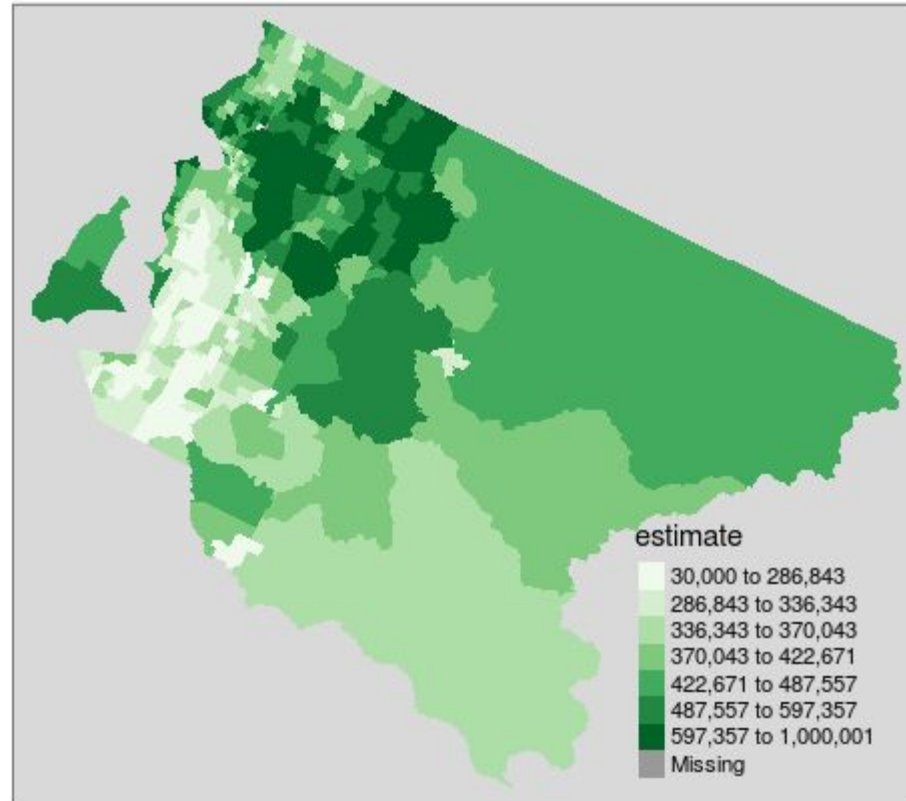
Zillow HVI and RI data



Predicting Median Housing Values in King County

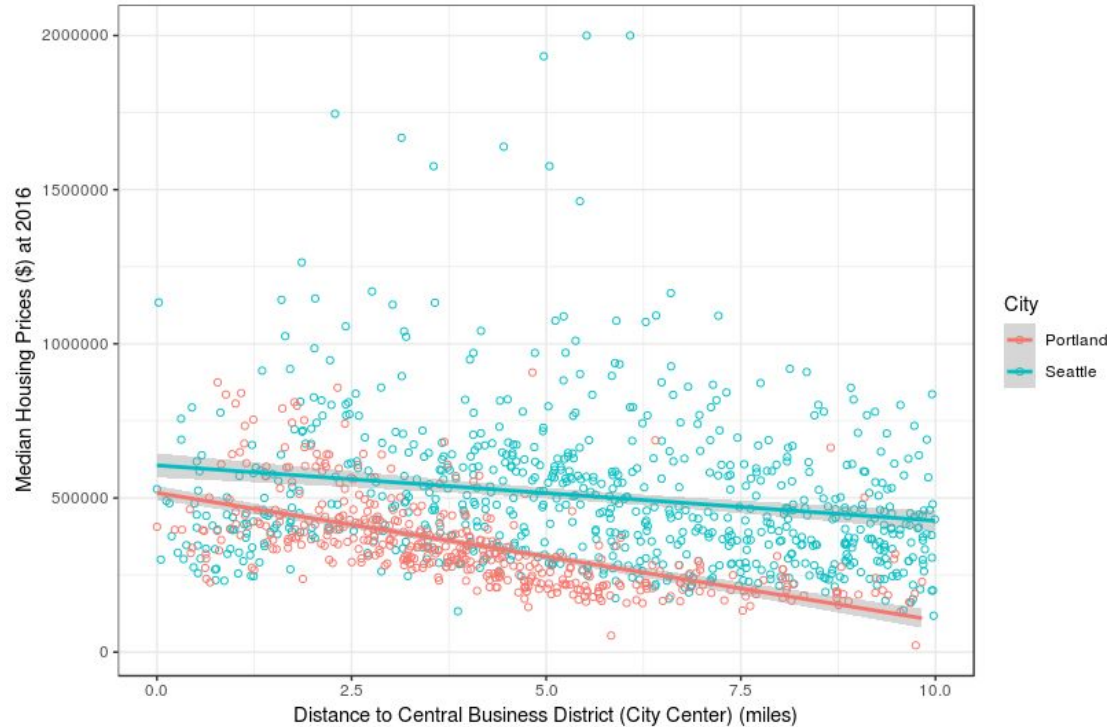
- 2016 5-year housing data for Washington State
- Discover contributing factors which may have an effect on the median housing values in king county
- Data Sourced from Census ACS Survey
 - Median Housing Value in King County
 - Unemployment Rate in King County
 - Median Household Income in King County
 - Poverty Rate in King County
 - Poverty thresholds were derived from the cost of a minimum food diet multiplied by three to account for other family expenses.
 - Renter Occupied Households in King County

Median Housing Values Across King County



Median housing price with distance to the city center of Seattle and Portland

(Spatial equilibrium in the Alons-Muth-Mills model)



Median housing prices will decline with distance to the city center. The speed at which they decline depends on transportation costs.

Least Square Model 1

y=median housing price values

x1=unemployment rate, x2=median household income, x3=poverty rate,

x4=renter occupied households

Call:

```
lm(formula = y ~ x1 + x2 + x3 + x4, data = mydata)
```

Residuals:

Min	1Q	Median	3Q	Max
-323267	-64256	-12866	46705	991447

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	322558.3781	20748.4930	15.546	< 0.0000000000000002	***
x1	-10033667.9510	476200.2731	-21.070	< 0.0000000000000002	***
x2	5.1053	0.0673	75.855	< 0.0000000000000002	***
x3	1179.0368	242.2251	4.868	0.00000116	***
x4	70.6241	2.5212	28.012	< 0.0000000000000002	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 105400 on 5905 degrees of freedom

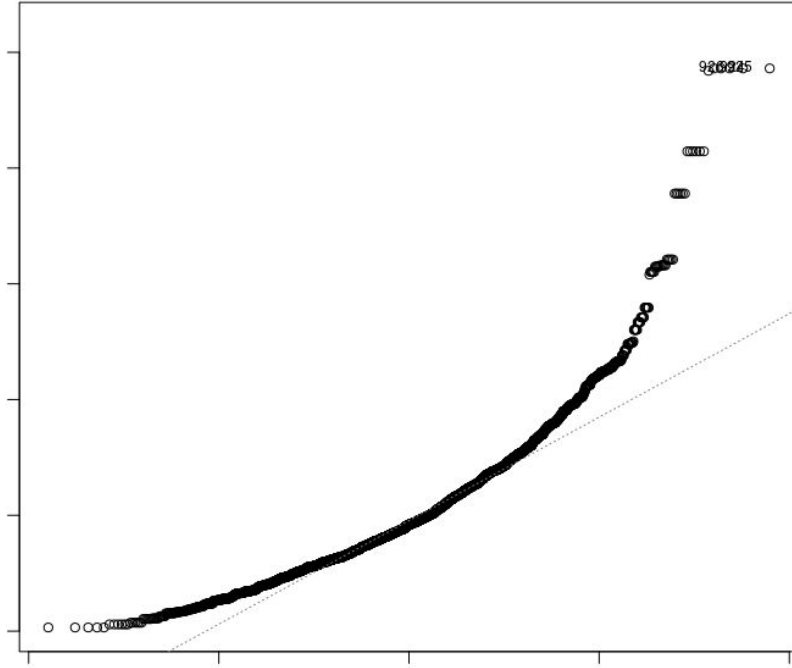
(864 observations deleted due to missingness)

Multiple R-squared: 0.6309, Adjusted R-squared: 0.6306

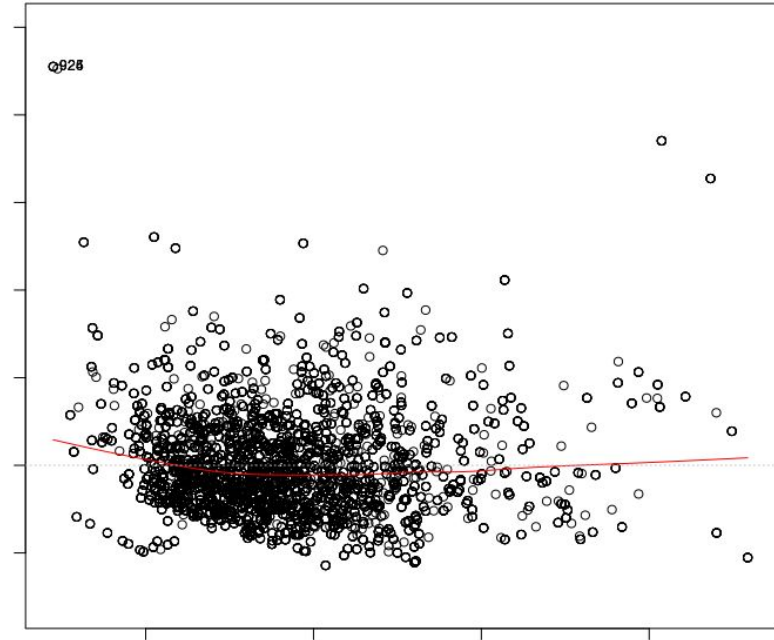
F-statistic: 2523 on 4 and 5905 DF, p-value: < 0.00000000000000022

Diagnostic plots: checks for heteroscedasticity, normality, and influential observations.

Normal Q-Q



Residuals vs Fitted



Least Square Model 2

y=median house pricing values property

x1=unemployment rate, x2=median household income, x3=poverty rate

Call:

```
lm(formula = y ~ x1 + x2 + x3, data = mydata)
```

Residuals:

Min	1Q	Median	3Q	Max
-253889	-72937	-12237	52154	900448

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	562468.49513	20113.10414	27.965	< 0.0000000000000002 ***
x1	-13660875.99651	487713.04294	-28.010	< 0.0000000000000002 ***
x2	4.39980	0.06643	66.237	< 0.0000000000000002 ***
x3	1940.09717	256.16822	7.574	0.0000000000000419 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 112200 on 5906 degrees of freedom

(864 observations deleted due to missingness)

Multiple R-squared: 0.5818, Adjusted R-squared: 0.5816

F-statistic: 2739 on 3 and 5906 DF, p-value: < 0.00000000000000022

Models Comparison and Subset selection

```
> AIC(fit1)
[1] 153486
> AIC(fit2)
[1] 154221.4
```

```
> anova(fit1, fit2)
Analysis of Variance Table
```

Model 1: $y \sim x_1 + x_2 + x_3 + x_4$

Model 2: $y \sim x_1 + x_2 + x_3$

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	5905	65630286915165				
2	5906	74351290248429	-1	-8721003333264	784.66	< 0.00000000000000022 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
> step$anova # display results
Stepwise Model Path
Analysis of Deviance Table
```

Initial Model:

$y \sim x_1 + x_2 + x_3 + x_4$

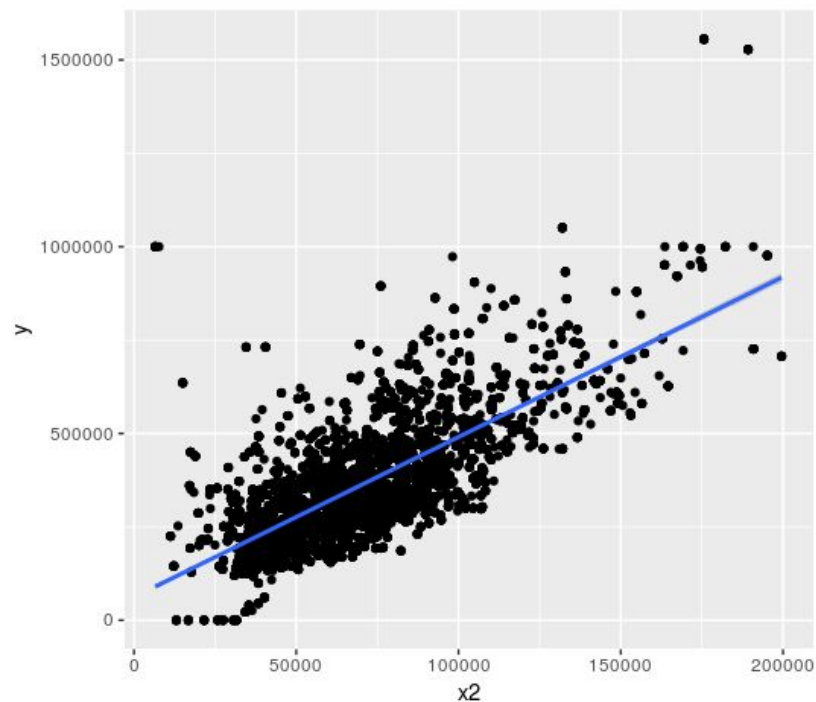
Final Model:

$y \sim x_1 + x_2 + x_3 + x_4$

Relative Importance

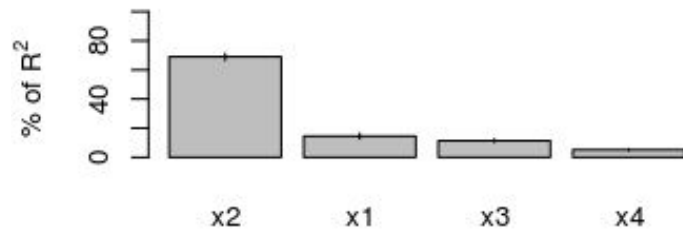
```
install.packages("relaimpo")  
library(relaimpo)  
calc.relimp(fit1,type=c("lmg","last","first","pratt"),  
            rela=TRUE)
```

```
# Bootstrap Measures of Relative Importance (1000 samples)  
boot <- boot.relimp(fit1, b = 1000, type = c("lmg",  
                                             "last", "first", "pratt"), rank = TRUE,  
                   diff = TRUE, rela = TRUE)  
booteval.relimp(boot) # print result  
plot(booteval.relimp(boot,sort=TRUE)) # plot result
```

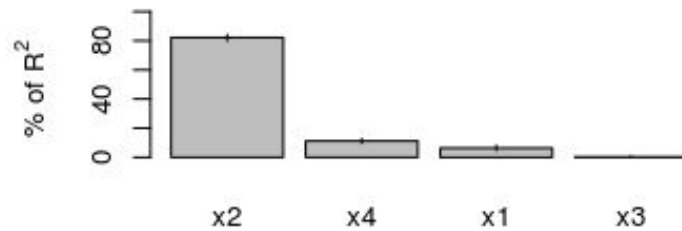


Relative importances for y
with 95% bootstrap confidence intervals

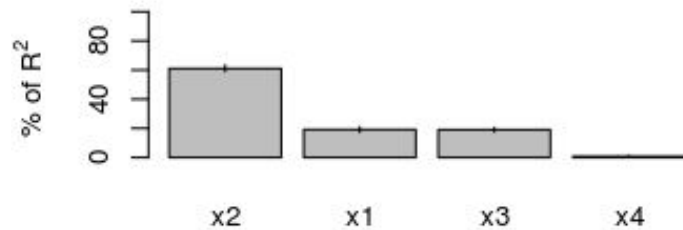
Method LMG



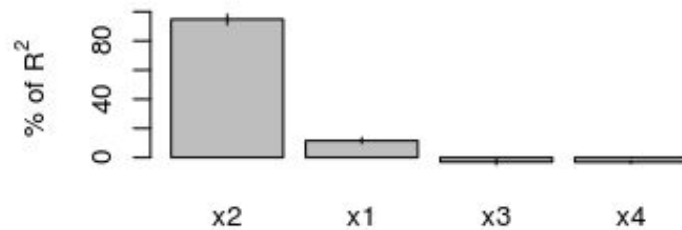
Method Last



Method First



Method Pratt



$R^2 = 63.09\%$, metrics are normalized to sum 100%.

References

- Census ACS Survey
 - <https://www.census.gov/programs-surveys/acs>
- Point-in-Time (PIT) estimate data
 - <https://www.hudexchange.info/resource/3031/pit-and-hic-data-since-2007/>
- Eviction Data
 - <https://evictionlab.org/>
- Zillow Home and Rent Indices
 - <https://www.zillow.com/research/data/>
- The economics of homelessness in Seattle and King County
 - <https://www.mckinsey.com/featured-insights/future-of-cities/the-economics-of-homelessness-in-seattle-and-king-county>

Thank you for your time!

