



Predicting Home Sale Prices in **West Chicago Neighborhoods**

Tara Ziegler

October 09, 2020

Sellers are seeking information about the **resale value** of their homes to make decisions.

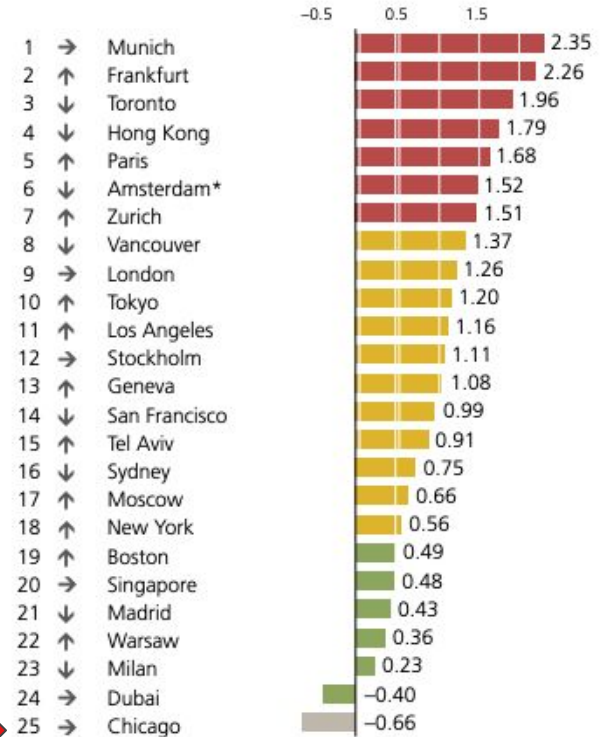


Chicago is the **most undervalued** market of 25 major international cities.

#25

UBS Global Real Estate Bubble Index

Index scores for the housing markets of select cities, 2020



■ Bubble risk (above 1.5) ■ Overvalued (0.5 to 1.5) ■ Fair-valued (-0.5 to 0.5) ■ Undervalued (-1.5 to -0.5) ↓ ↑ ↔ Change vs. 2019

Source: UBS * Index altered due to data source revision. For explanation see the section on Methodology & data on page 28.





WICKER
PARK



The data set is sourced from Zillow's website.

Data: 3 years of “Sold” listings scraped from Zillow.com

Zip Codes Selected: West Town, West Loop, Bucktown, Wicker Park

Target: Home Sale Price



Various tools were used to support this project.



BeautifulSoup



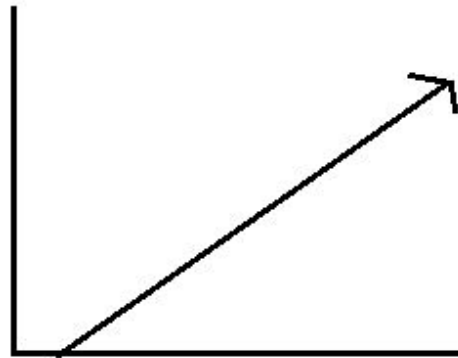
Selenium

Preprocessing scraped data improves the model's accuracy.

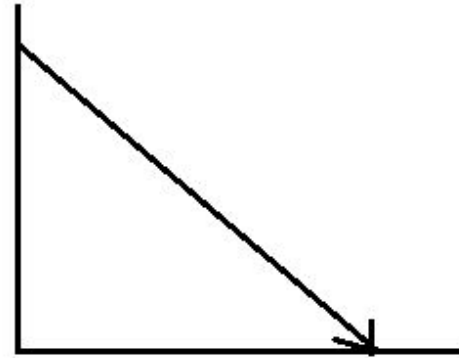
- Price **unavailable**
- Bedroom/Bathroom unavailable
- Price > \$10M
- Less than \$50k
- Zillow Listing Errors (ex: 26k sq foot home)
- Listing Type: Vacant Land, N/A, Other (Commercial)



The modeling technique used is called “**Linear Regression**”.



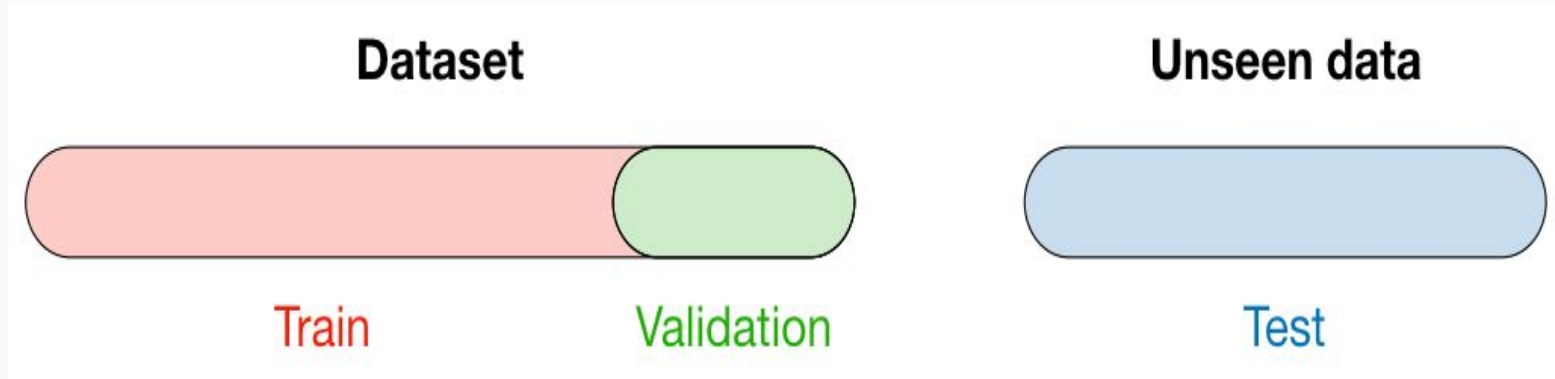
Positive Linear Relationship



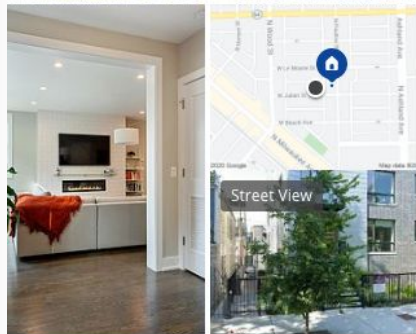
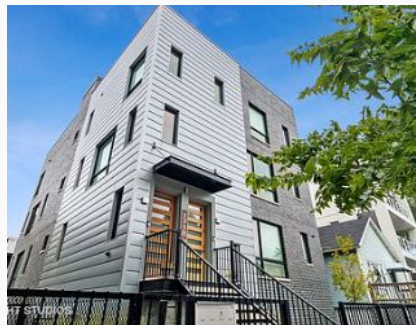
Negative Linear Relationship



Train, Validate, & Test!



Over 1,000 Zillow listings were **missing** square footage data.



Zillow

[Edit](#) [Save](#) [Share](#)

3 bd | 3 ba | -- sqft

1718 W Julian St UNIT 1N, Chicago, IL 60622

● Sold: \$640,000 | Sold on 09/24/20 | Zestimate®: \$659,09

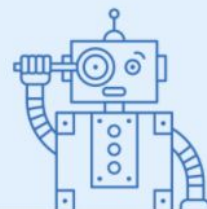
Est. refi payment: \$4,071/mo [Get current rates](#)

[Home value](#) [Owner tools](#) [Home details](#) [Neighborhood details](#) [Similar](#)

Home value

It looks like this property has **missing facts**, which can affect the accuracy of home value estimates.

Is this your home? [Claim it](#) and update home facts!



Zestimate
\$659,092



$$\frac{\text{Average Price (available using training data)}}{\text{Average Sq Ft (available using training data)}} = \text{Average Price/Sqft}$$

$$\text{Average Price/Sqft} \times \text{Listing Price (train \& test)} = \text{Estimated Sq Ft (input for missing data in train \& test)}$$



Results



1: The model scored how **square footage** impacts price.

Training Data: 0.963

Validation Data: 0.934



2: The model scored how **additional features** impacts home sale price.

- Square Footage
- Listing Type
- # of Bedrooms
- # of Bathrooms
- Parking (Y/N)



2: The model scored how **additional features** impacts home sale price.

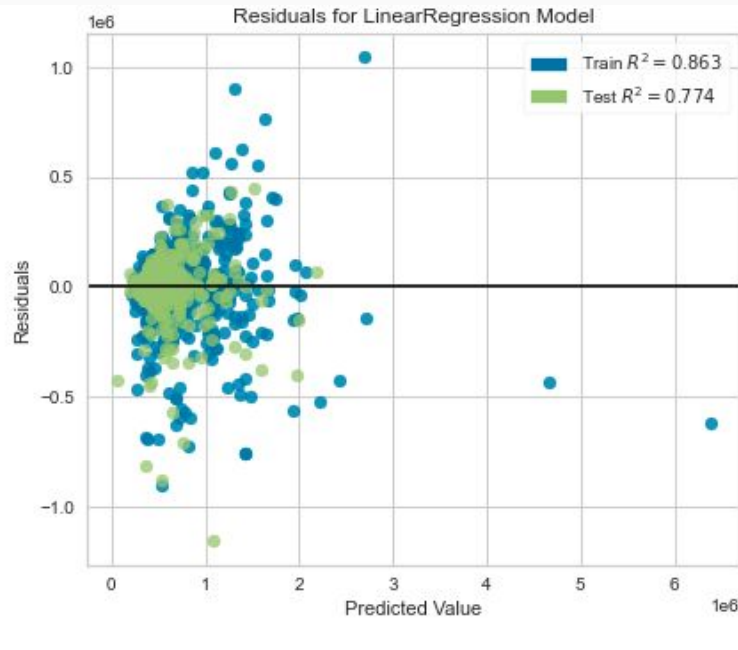
Training Data: 0.765

Validation Data: 0.863

Test Data: 0.774



Residuals for Linear Regression Model



Conclusion & Future Work

- Square footage can have a profound effect on a home's sale price.
- Seasonality of Home Sale Transactions
- Exterior Features (Roofdeck, Patio)
- HOA fees
- Year built
- Additional Machine Learning Techniques



Questions?



Appendix

Estimating Square Footage for Missing Data:

- Actual versus Predicted square footage for Train data is \$427.
- Actual versus Predicted square footage for Train data is \$424.

UBS International Bubble Index Report

- <https://www.ubs.com/content/dam/static/noindex/wealth-management/ubs-global-real-estate-bubble-index-en-1509323.pdf?campID=UC:E:601227:601241:269449152:0:543958130:543958132:en:543272033:::>



Predicted Sale Prices vs Actual Sale Prices

