

# Redfin Analysis Presentation

Regression Analysis on homes on Redfin in a variety of cities in Illinois .

Snizhana Kurylyuk

*Metis Data Science and Machine Learning Student*



A client reached who is thinking of buying a single family home and doesn't know which area is best. They wanted me to build them a regression to allow them to not overpay for a home and provide analysis on the types of homes they would find in each neighborhood.

## Objective

Web Scrape Redfin for various home features and analyze the home prices based on those features. Then build a predictive regressive model to predict the price of a home.

## Design

I focused homes are currently on the market.

### Wilmette Homes for Sale

[See Market Insights](#)

Price

No min

to

No max

[More Filters](#) ▾41 of 69 homes · Sort: [Recommended](#) ▾[Photos](#)[Table](#)**\$129,700**

2 Beds 1 Bath 934 Sq. Ft.

400 Laurel Ave Unit 2N, Wilmette, IL 60091

**\$2,250,000**

7 Beds 5.5 Baths — Sq. Ft.

1110 Lake Ave, Wilmette, IL 60091

**\$1,995,000**

7 Beds 6.5 Baths — Sq. Ft.

111 Robsart Rd, Kenilworth, IL 60043

**\$1,199,000**

5 Beds 5 Baths 4,040 Sq. Ft.

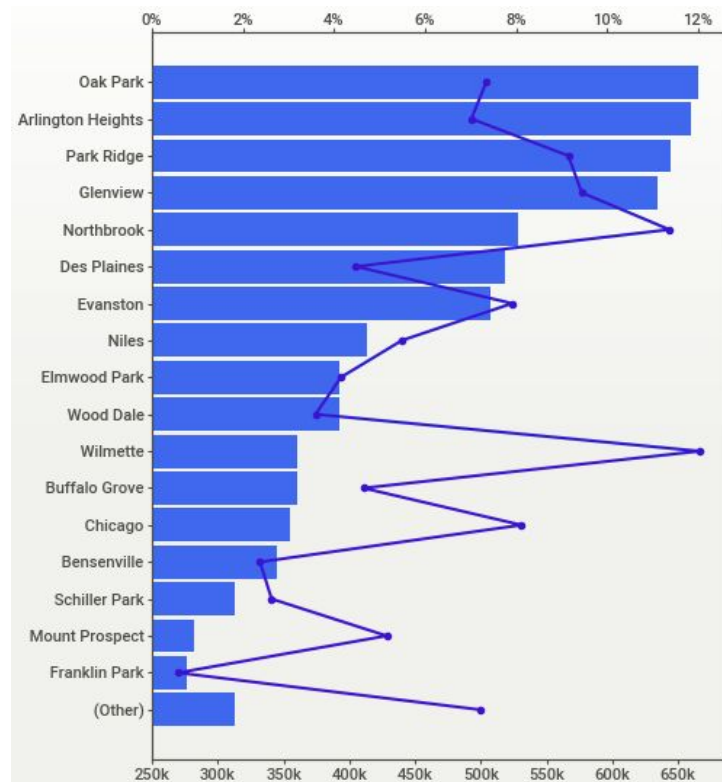
430 Lake Ave, Wilmette, IL 60091

# Data

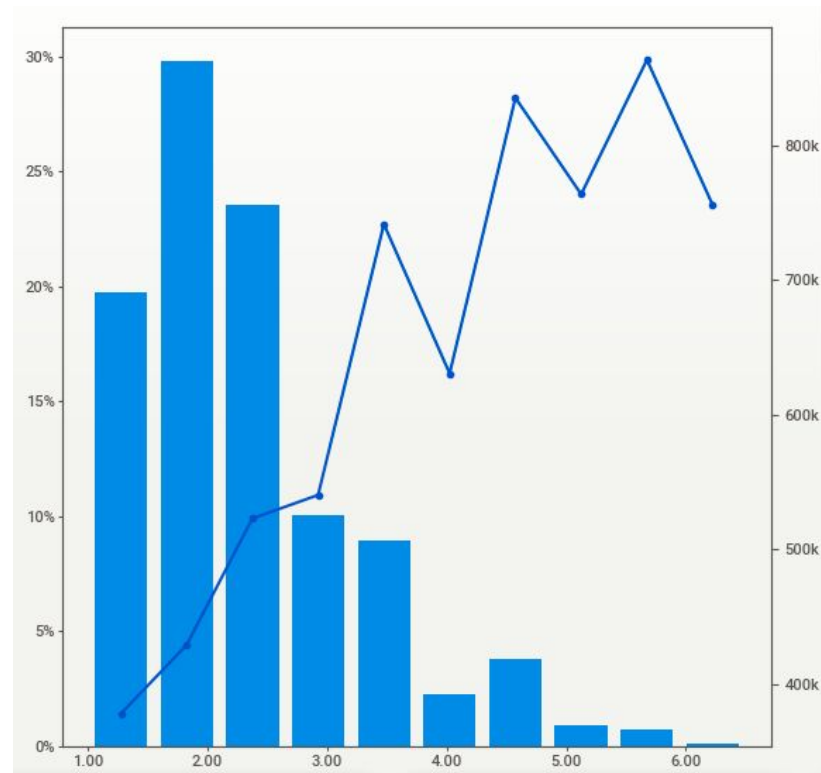
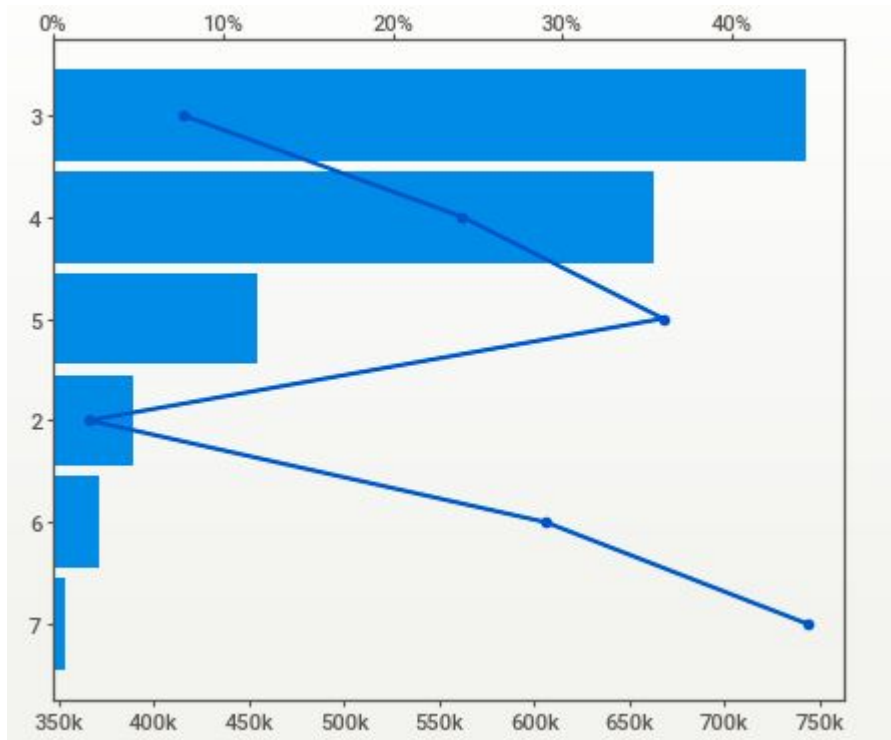
Single family homes currently for sale in  
in *Chicago, Park Ridge, Niles, Elmwood  
Park, Des Plaines, Mount Prospect,  
Northbrook, Highland Park, Glenview,  
Wood Dale, Bensenville, Addison, Schiller  
Park, Norridge, Franklin Park, Arlington  
Heights, Buffalo Grove, Morton Grove, Oak  
Park, Evanston, Wilmette, Northfield.*

## Home Features

- Price/Sq.Ft
- Lot Size
- Year Built
- Tax Annual Amount
- Tax Year
- Tax Exemptions
- Bedrooms
- Bathrooms
- Sq\_Footage
- Price
- City
- Zip Code

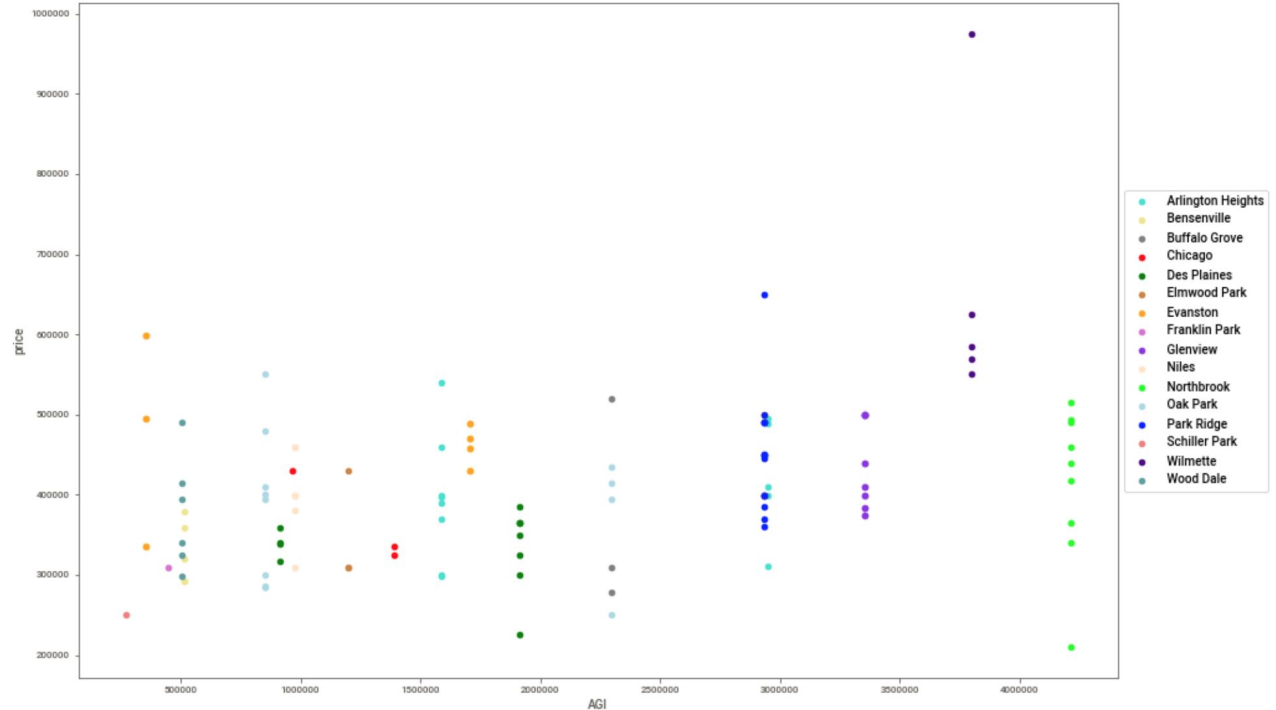


# Bedrooms and Bathrooms Effect on Price

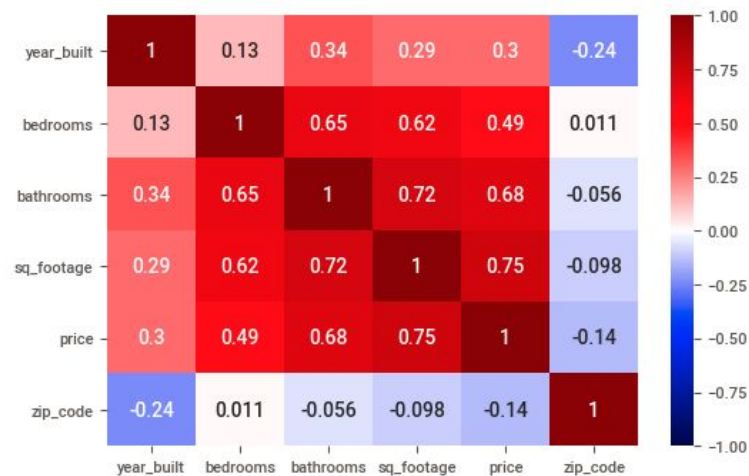
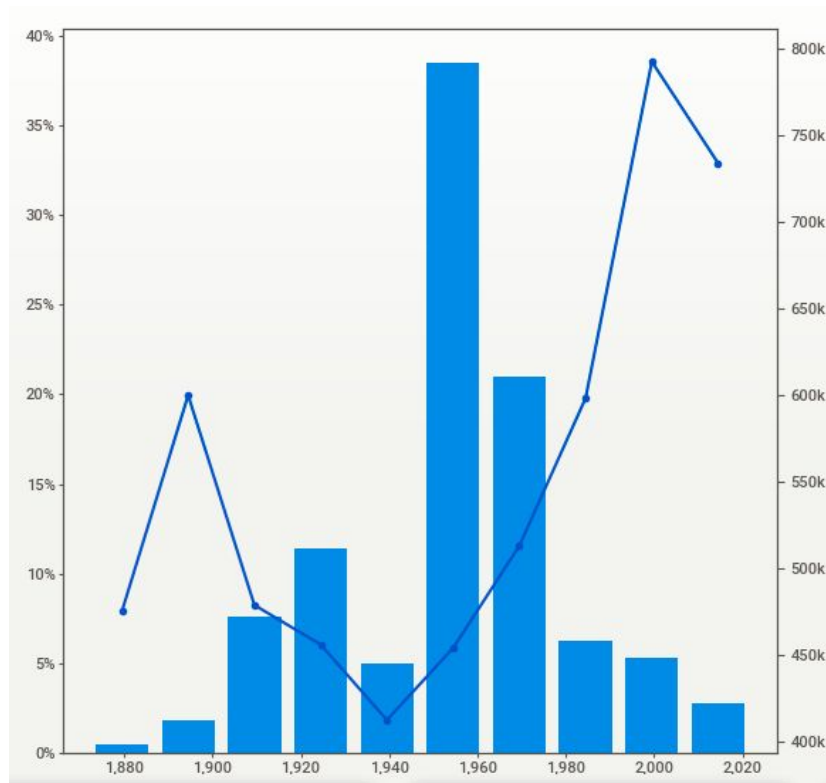


# IRS Income Data

Prices of 3 Bedroom 2 bath homes based on City and Adjusted Gross Income.

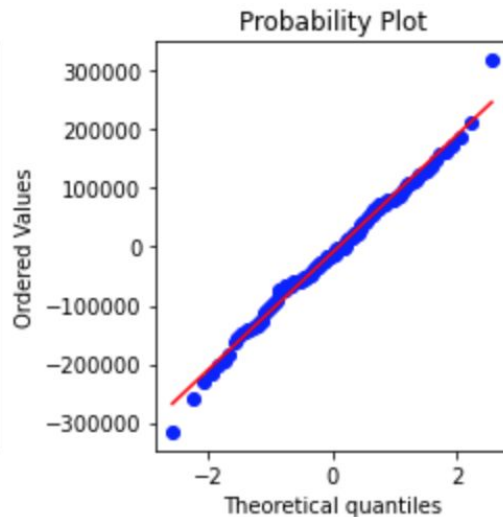
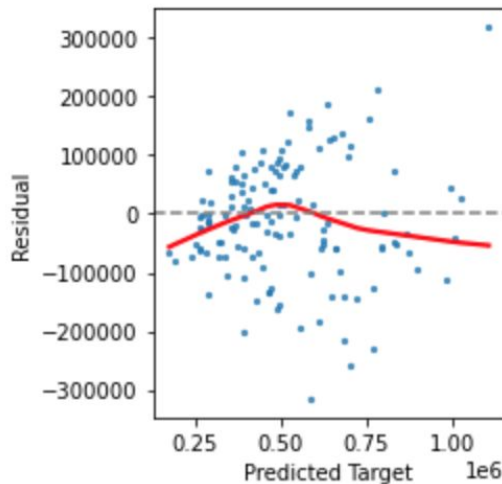
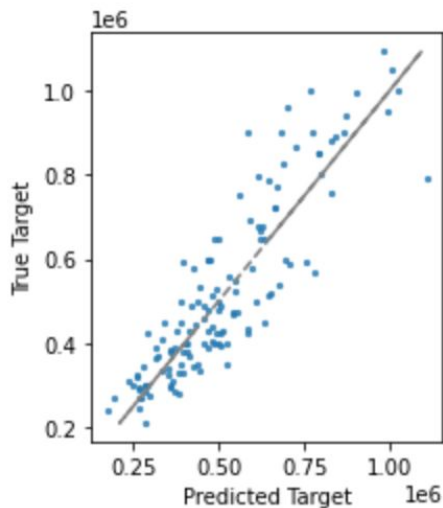


# Home Price By Sq Footage



# Linear/Ridge/Lasso Regressions

Diagnostic Plots

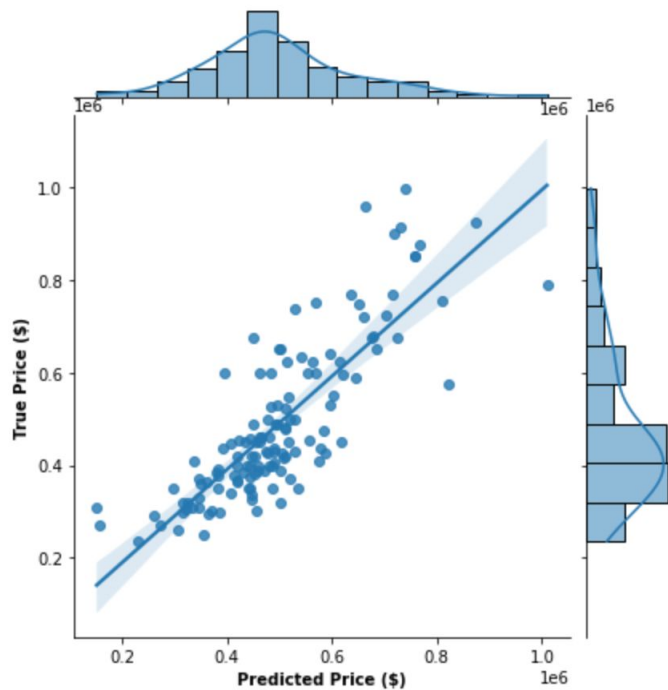


Ridge Model - Reduces effects  
of collinear features

$R^2$  Adjusted = .72

MAE = 71,983

# Prediction



city	mean_variance
Wilmette	-168368.501831
Franklin Park	-135686.593926
Elmwood Park	-32866.849522
Evanston	-18825.792418
Bensenville	-11936.178342
Glenview	-11247.588059
Niles	2237.109078
Park Ridge	7701.089465
Buffalo Grove	8224.306367
Oak Park	15557.011223
Des Plaines	21935.746058
Wood Dale	26778.537069
Schiller Park	38426.588287
Arlington Heights	39070.686009
Chicago	44376.887222
Northbrook	53761.637344

city	AGI
Northfield	6279394.0
Highland Park	5423212.0
Northbrook	4212044.0
Wilmette	3799640.0
Glenview	3356144.0
Park Ridge	2936448.0
Buffalo Grove	2294931.0
Mount Prospect	2228885.0
Des Plaines	1911734.0
Evanston	1705947.0
Arlington Heights	1587983.0
Chicago	1390194.0
Elmwood Park	1200849.0
Addison	1115774.0
Niles	977564.0
Morton Grove	963859.0
Oak Park	850044.0
Norridge	774435.0
Bensenville	516331.0
Wood Dale	506236.0
Franklin Park	448028.0
Schiller Park	270330.0



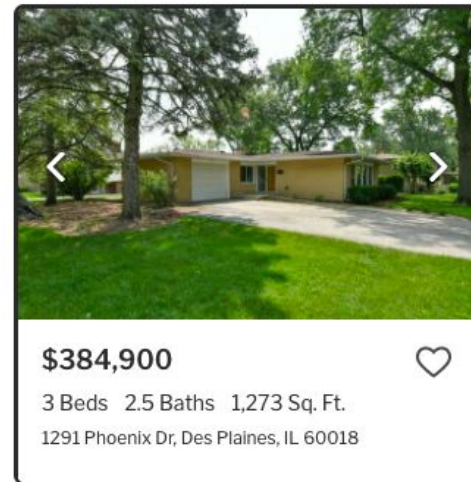
# Prediction Comparison



My Prediction: \$495,505



My Prediction: \$457,352



My Prediction: \$381,013

Questions?