Rental Property Capstone

Kristian Knudtson

Prospecting the U.S. Rental Market “Mining for cash flow”

Motivation and Introduction

- Post 2008 Financial Crisis, large investors entered the rental property market in a big way with a reported 1M+ people now living in homes owned by large landlords (see Invitation Homes).
- Many of these institutional investors are buying these properties based on models, leveraging large datasets (see Zillow, Opendoor).
- With transparency increasing, regular investors can take a smarter approach to finding real estate opportunities as well.

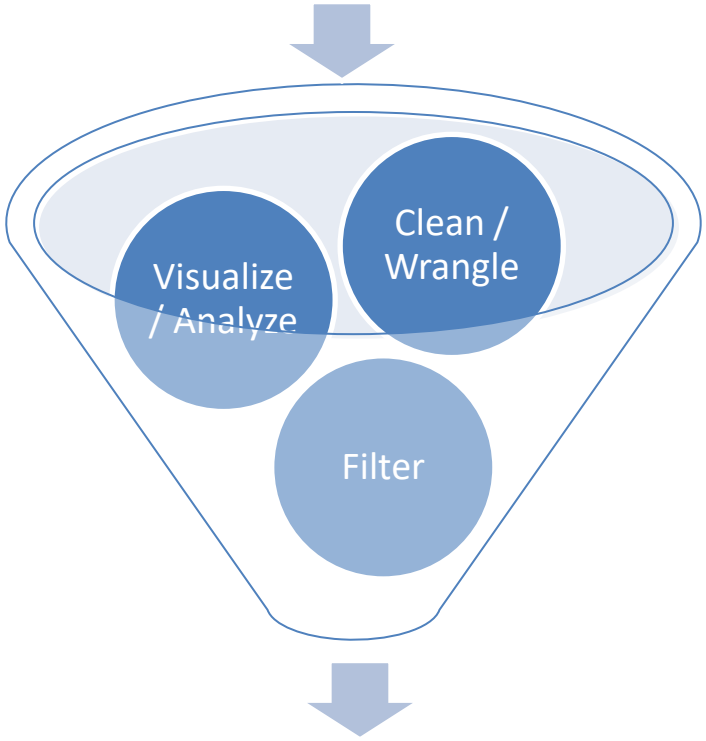
Problem Statement and Approach

- **Problem:** Regular property investors are unable to compare returns across vast regions or have little flexibility in their assumptions; paving the wave for costly outcomes.

- **Approach:** A scalable and flexible rental property machine learning model that can look at data across U.S. neighborhoods.

There are a lot of neighborhoods to look at!

4,691 Neighborhoods



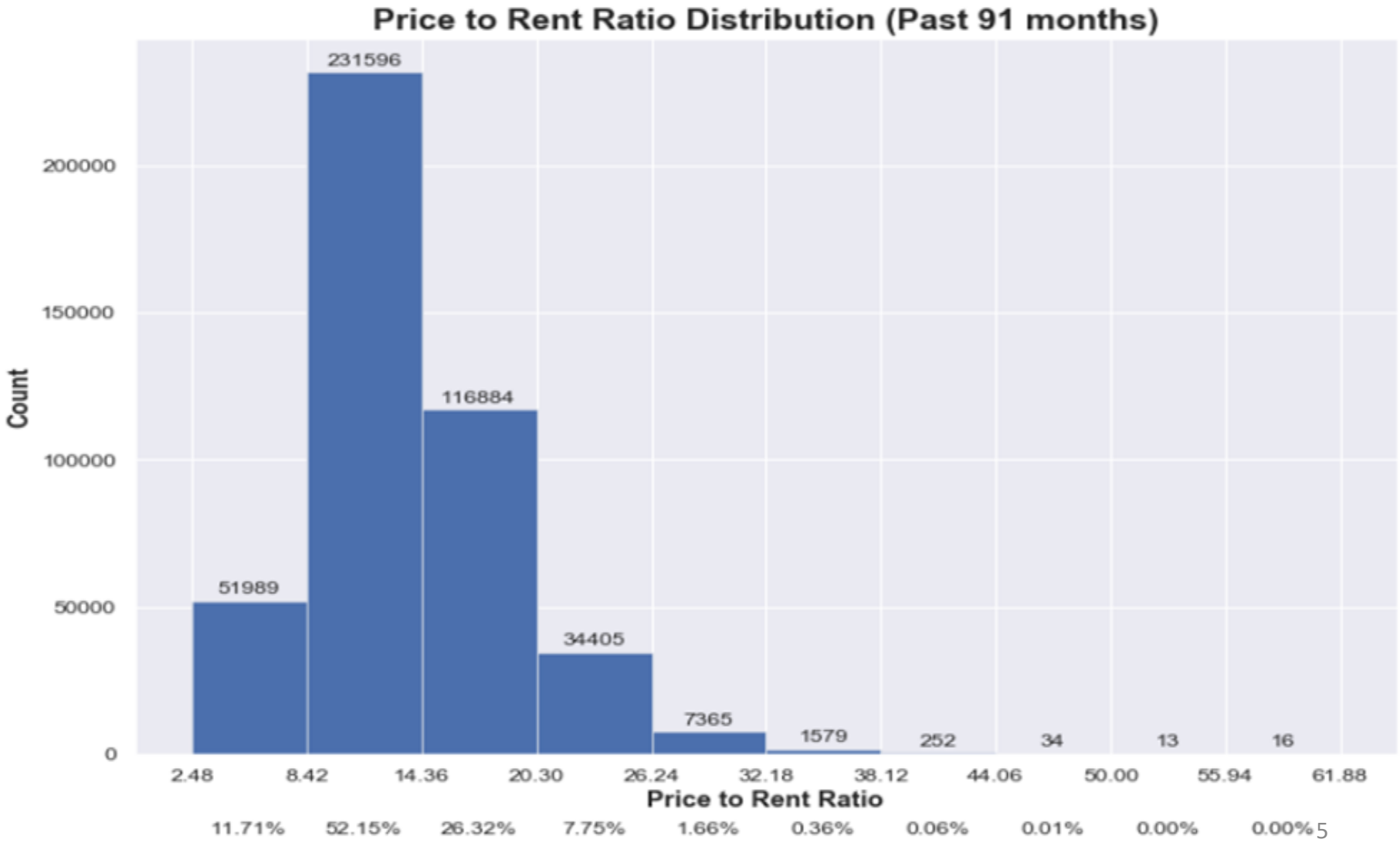
3,202 Neighborhoods.

Too many for a human to analyze, easy for a machine.

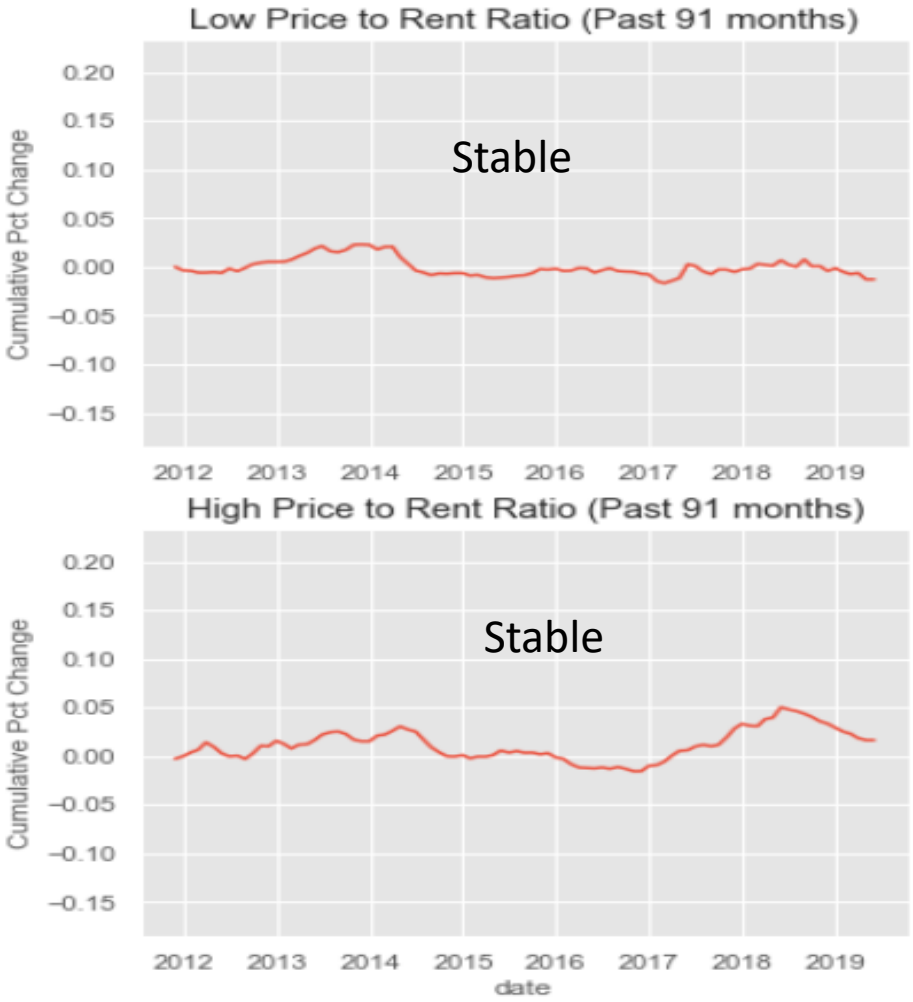
Filters:

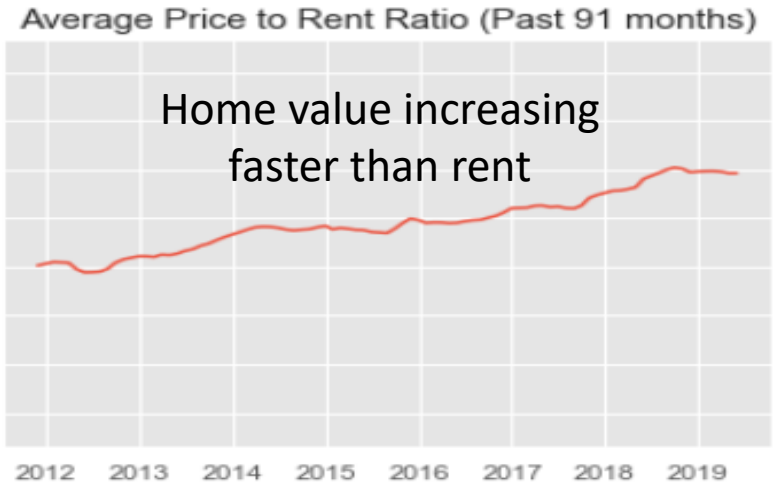
1) Focused on Single-Family Homes (SFH) under $500,000

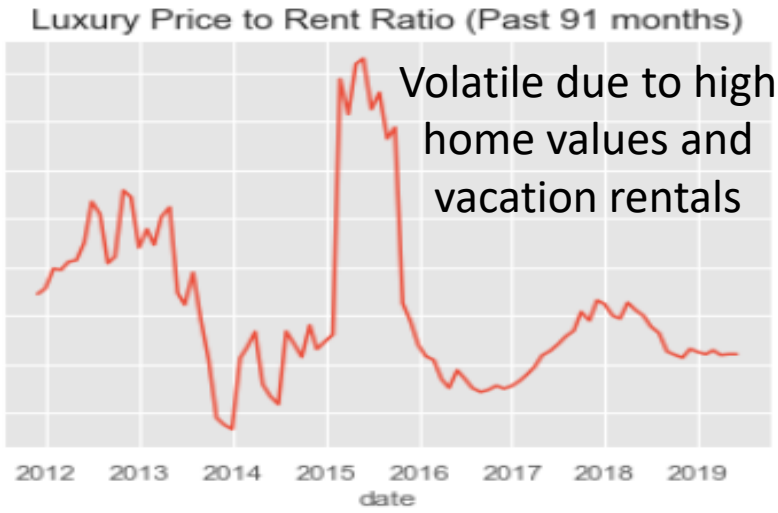
Wide differences in neighborhood price-to-rent ratios indicate opportunities are likely



Digging deeper into price-to-rent ratios segments illustrates different growth trends (number?)







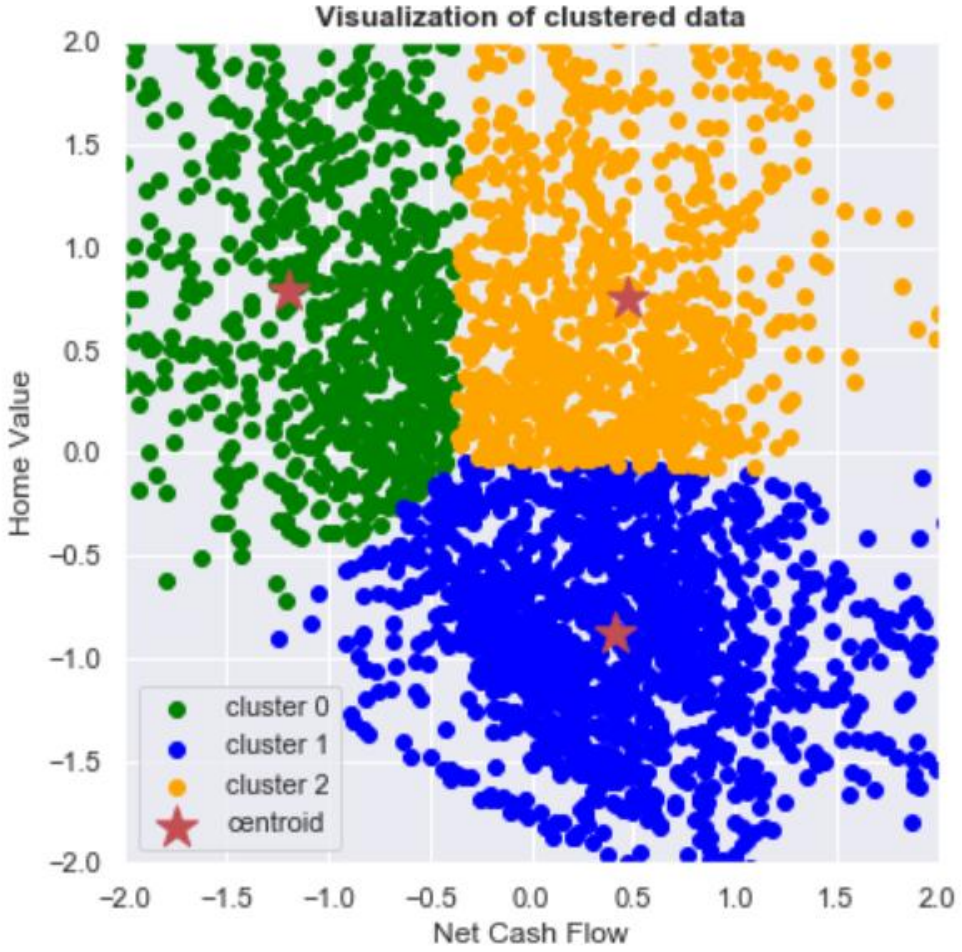
Clustering Approach

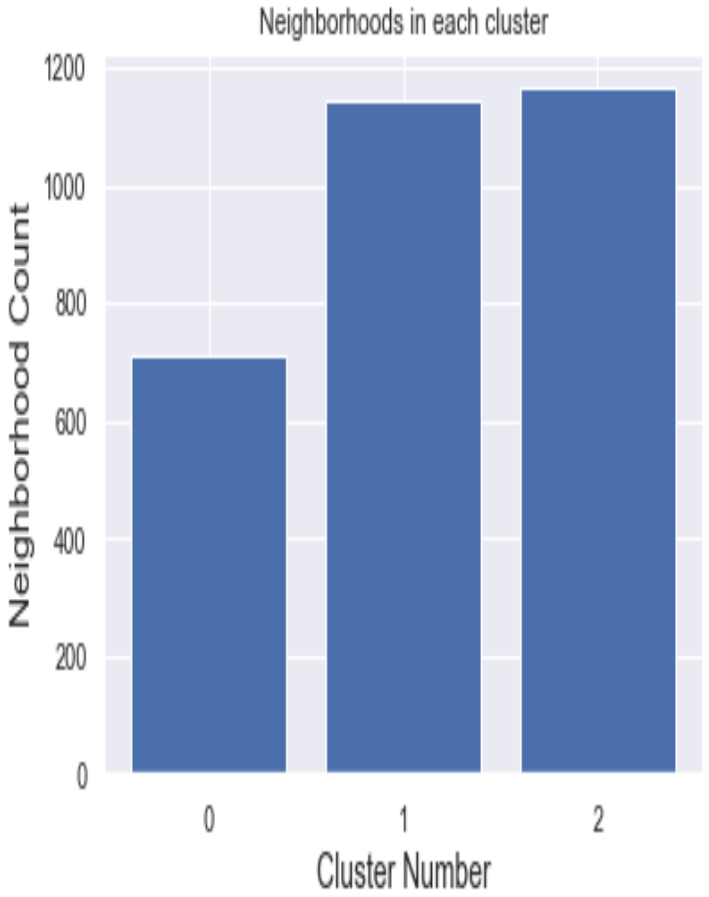
- Not predicting values or labels
- Focused on groupings that could inform investment strategy

Features
1) Neighborhood Estimated Net Cash Flow
2) Neighborhood Zillow Home Value Estimate - Single Family Residence

Model Results

- KMeans had best silhouette scores of three models tested
- n_clusters = 3 @ 0.38 silhouette score (weak structure)





Real World Application Approach

1) Subset a state in dataset where you have lived and sort on highest net cash flow neighborhoods. Look for cluster pattern.
2) Compare against all other states for relative attractiveness.

RegionName	State	net_cf	cluster
East Isles	MN	1242.0	0
Whittier	MN	727.0	1
Southeast Como	MN	589.0	2
Hawthorne	MN	456.0	1
Mckinley	MN	439.0	1
Willard Hay	MN	417.0	1
Marcy Holmes	MN	411.0	2
Folwell	MN	409.0	1
East Phillips	MN	405.0	1
Webber-Camden	MN	392.0	1

RegionName	State	net_cf	cluster
East Isles	MN	1242.0	0
Heritage Hills	NY	955.0	0
Indian Creek	CO	935.0	2
Brookwood	GA	877.0	2
Meadow Hills	CO	798.0	2
Cadwalader & Hillcrest	NJ	782.0	1
Chambersburg	NJ	777.0	1
Wynnefield	PA	767.0	1
Sable Ridge	CO	759.0	2
Bolton	GA	754.0	2

Expensive neighborhoods with many condos that carry high HOA fees, which aren't accounted for in the model!

Real World Application - Takeaways

- Clusters 0 and 2 have higher home values on average. This coincided with a higher prevalence of condos that tend to carry high HOA fees. Remove outliers?

- Cluster 1 is likely to produce more accurate results based on initial spot checks.

Conclusion

The Good

- Provides some clarity amongst 4,691 potential neighborhoods to invest in.
- Flexibility of model leaves room for various lenses on data
- Cluster 1 provided solid initial leads

The Bad

- Lack of property-level data is a disadvantage.
- Cluster shapes have no separation indicating model could benefit from additional features (crime, school ratings, etc.)

Appendix: Data Sources

Source	Data	Periodicity
Zillow Data	Median Home Values Median Rent Values Price-to-Rent Ratios 30 Year Interest Rates Size Rank (Neighborhoods)	Monthly October 2011 to Present
Value Penguin	Insurance Costs by State	2019 (Yearly)
U.S. Census Bureau, NAHB, American Community Survey	Property Tax Data for All U.S. Counties	As of 2014

Data Coverage:

45 States → 123 metros → 174 counties → 266 cities → 4,691 Neighborhoods

Investors Care About Cash Flow!

Net Cash Flow			
Component		**Assumption**	**Notes**
	Gross Rent	Zillow Estimate	
-	**Vacancy (months)**	One month each year	Flexible
=	**Gross Rent**		
-	**Management Fees**	8% of monthly rent	Flexible
-	**Operating Expenses**	- $250/month (Repairs = $200, Maintenance = $50) - Adjusted by value (* 250)	Flexible
-	**Taxes**	(Zillow Estimate * Multiplier @ 80%) * (County Avg/12)	Flexible multiplier
-	**Interest Payment**	- Zillow estimate on given date = purchase price - 30 year interest rates (equal to monthly average) - 20% down payment	- Function creates amortization table for each region given purchase date - Flexible down payment amount
-	**Insurance**	(Region Med. Value / Statewide Avg. Value) * Statewide average home insurance cost	Update Insurance data each year
=	**Net Cash Flow**		