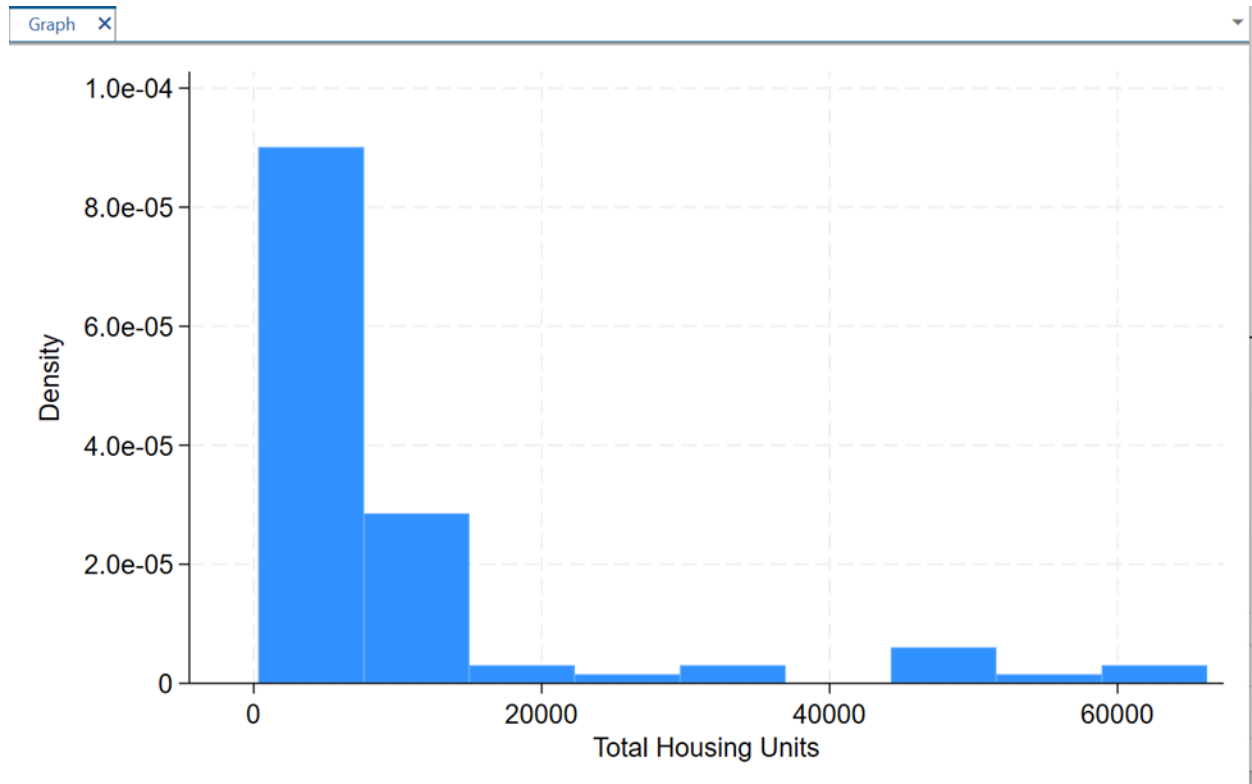


Seattle Housing Crisis Analysis

Part 2: Caylin Seitz & Christoph

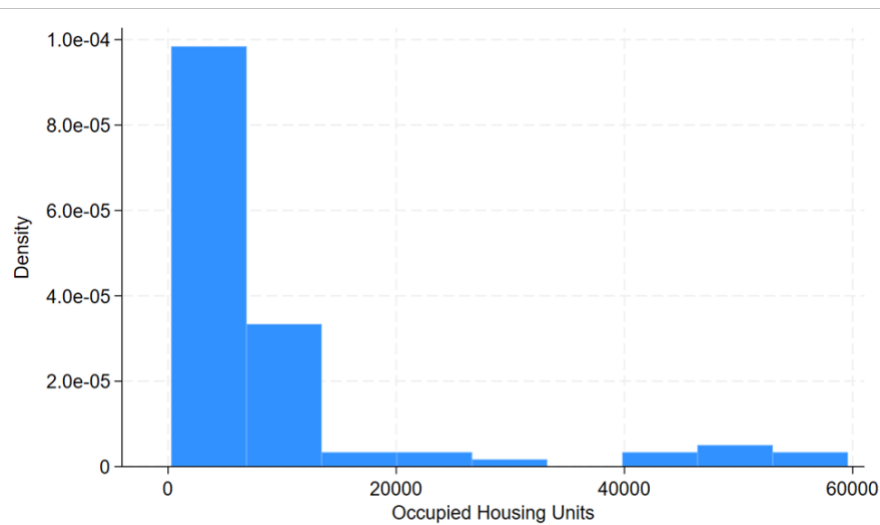
Univariate

Distribution of Total Housing Units in Seattle neighborhoods combined



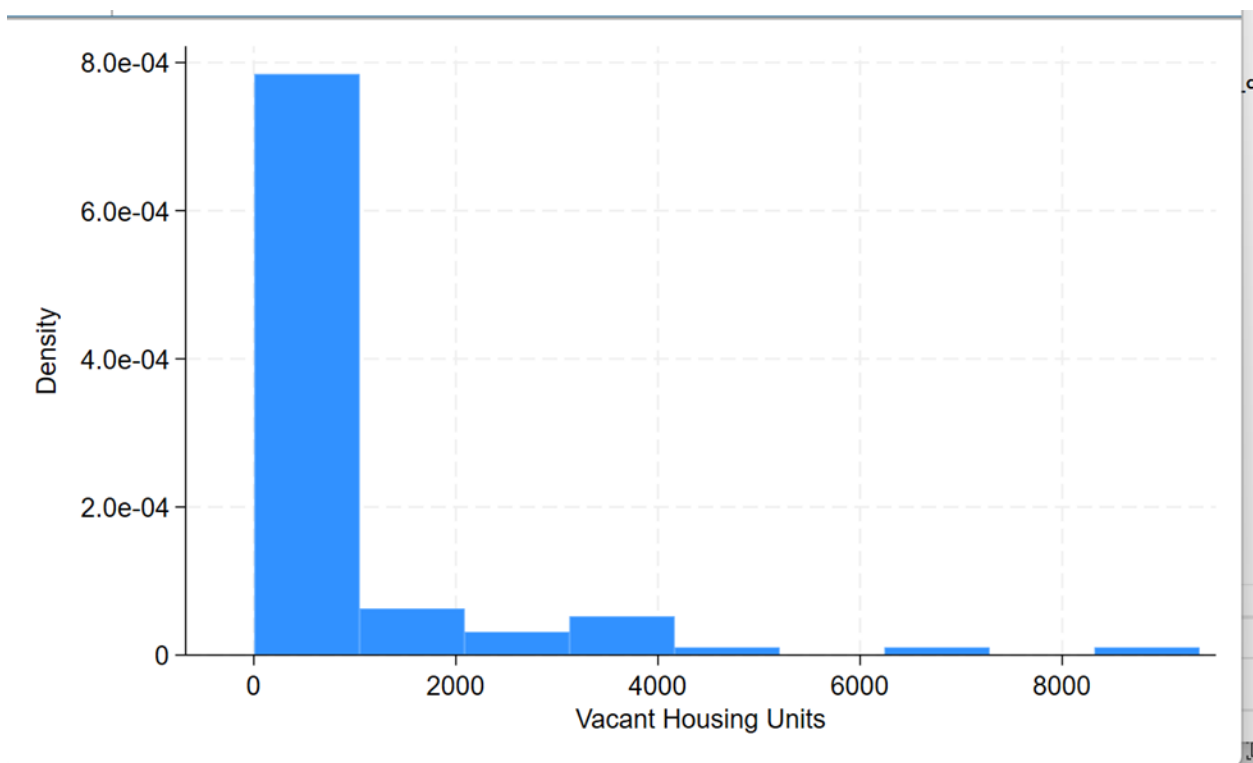
To understand Seattle's housing crisis, it is essential to observe the distribution of homes in Seattle. Comparing the histogram of the total housing units with the following histograms showing the distributions of occupied and vacant homes enables us to make initial comparisons and find gaps to investigate further.

Distribution of Occupied Housing Units:



The overall distribution of occupied housing units follows a similar distribution to the total housing units. This is expected as the occupied housing units combined with the vacant housing units amount to the sum of the total housing units.

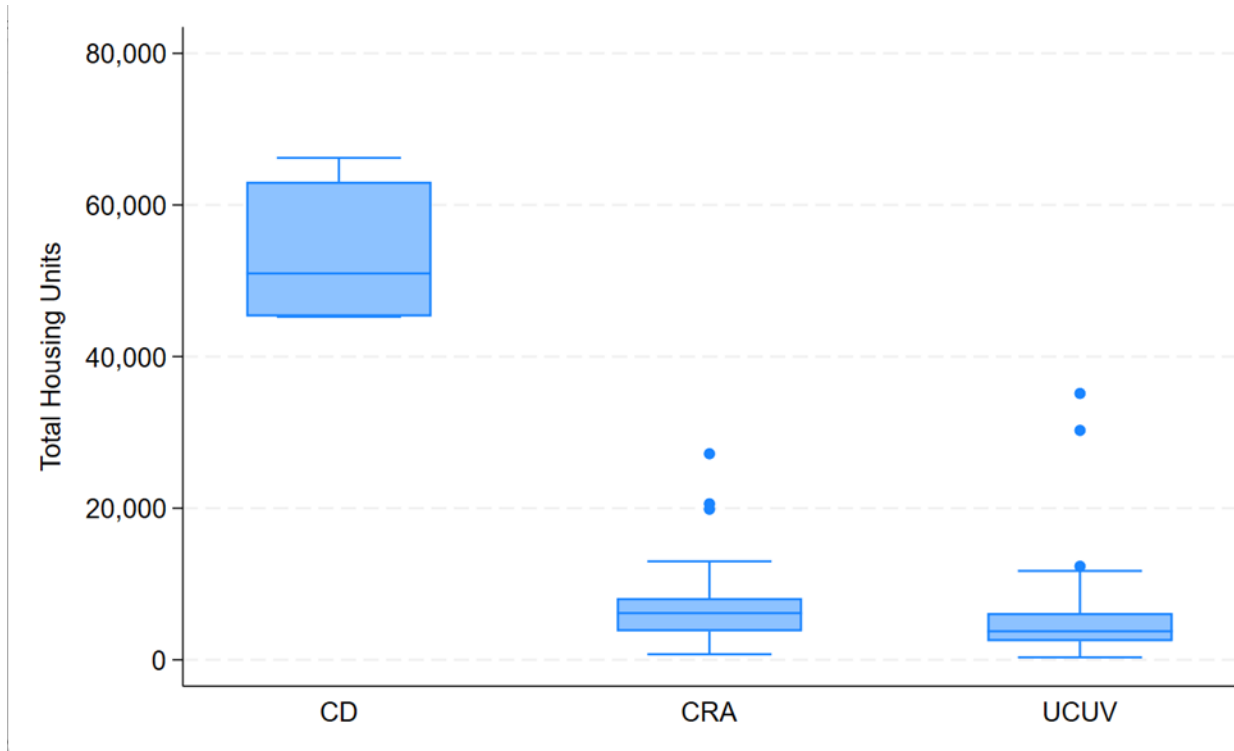
Distribution of Vacant Housing Units



Comparing the distribution of the vacant housing units to the occupied housing units, there are some interesting insights that can be derived. We can see that there is a significant amount of vacant housing units, which raises concern about the supply of available housing units and the possible relation of unaffordability.

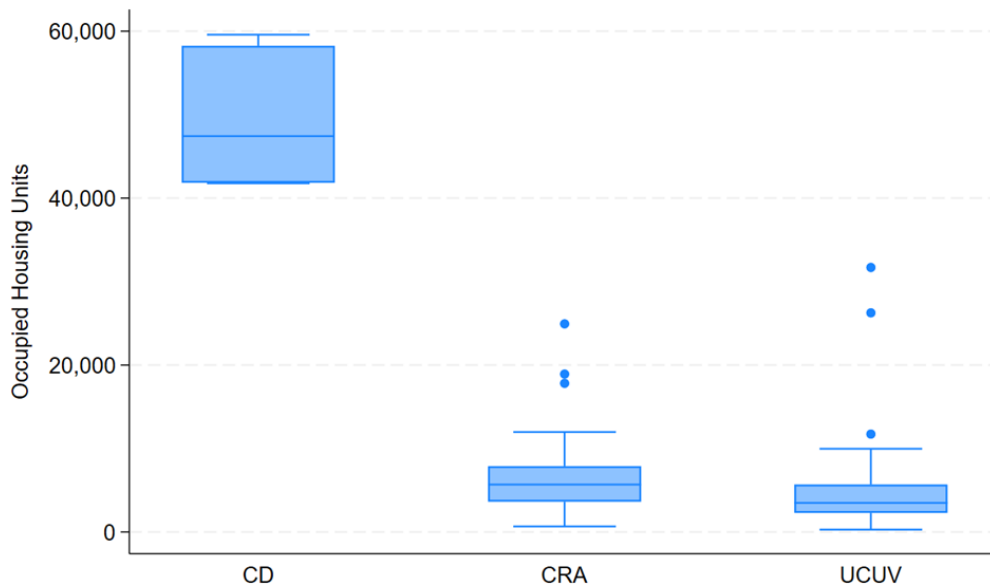
Bivariate:

Total housing units by neighborhood type

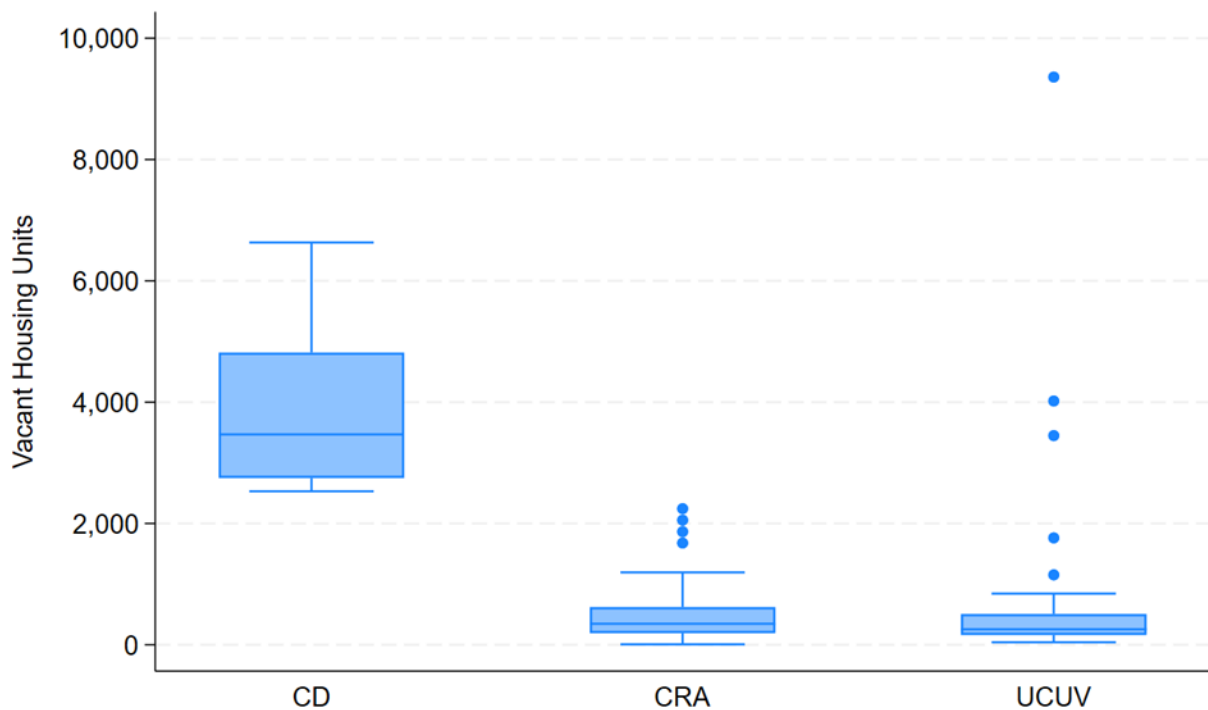


Disaggregating the total housing units by neighborhood type, we can compare the three main categories in the dataset for City of Seattle Council Districts, Comprehensive Plan Growth Areas, and Community Reporting Areas. These boxplots further detail the distribution of all housing by subcategory.

Occupied units by neighborhood type



Vacant housing by neighborhood type. This plot shows that the highest occupied housing units are in the CD (Council District) neighborhoods. CRA (Community Reporting Areas) and UCUV (Comprehensive Plan Growth Areas) show an almost identical distribution.



The boxplots of Vacant Housing shows a significantly smaller range between the three neighborhoods. While CD neighborhoods are still much greater for both Vacant and Occupied, the gap between CD and the other neighborhoods is smaller compared to the Occupied plot. This helps to describe areas of focus for determining target areas of the housing crisis.

Descriptive Tables: Basic statistics further describing the total, vacant, and occupied housing in the dataset

Total Housing Units				
	Percentiles	Smallest		
1%	328	328		
5%	1485	736		
10%	2298	1460	Obs	92
25%	3385.5	1485	Sum of wgt.	92
50%	5518.5		Mean	12154.43
		Largest	Std. dev.	22349.21
75%	9275.5	51843		
90%	30267	63064	Variance	4.99e+08
95%	50967	66214	Skewness	5.096399
99%	178765	178765	Kurtosis	35.43884

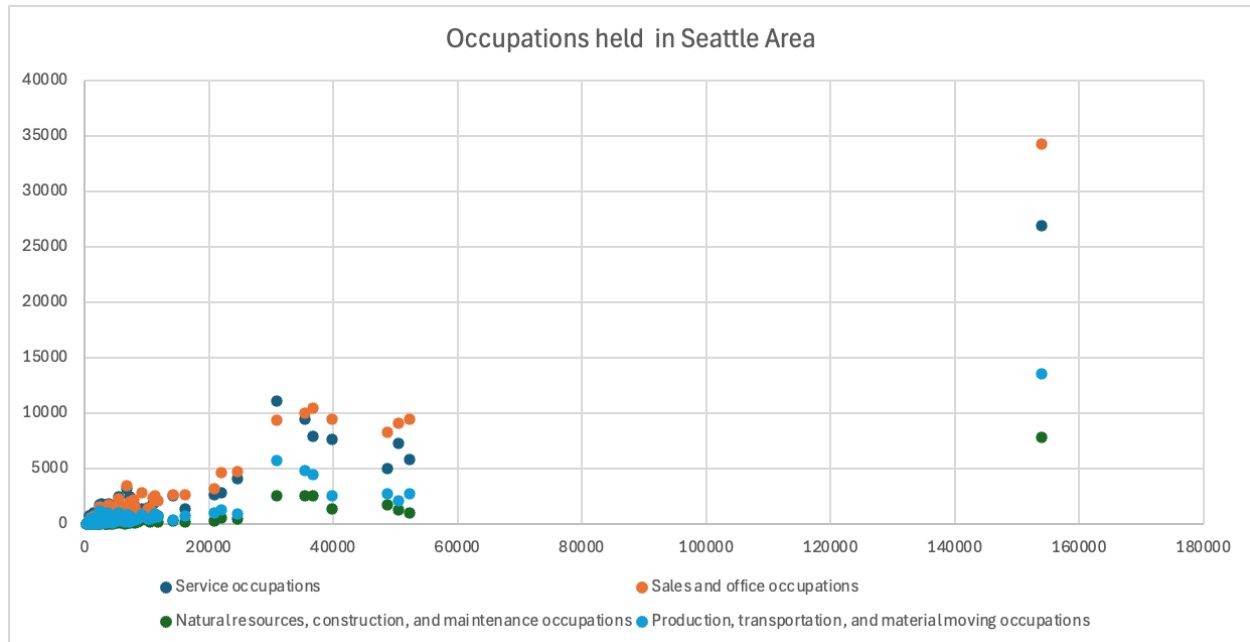
.

Occupied Housing Units				
	Percentiles	Smallest		
1%	283	283		
5%	1444	662		
10%	2105	1346	Obs	92
25%	3130	1444	Sum of wgt.	92
50%	5149		Mean	11267.8
		Largest	Std. dev.	20974.71
75%	8546.5	49313		
90%	26249	58248	Variance	4.40e+08
95%	47425	59582	Skewness	5.225018
99%	169406	169406	Kurtosis	36.93512

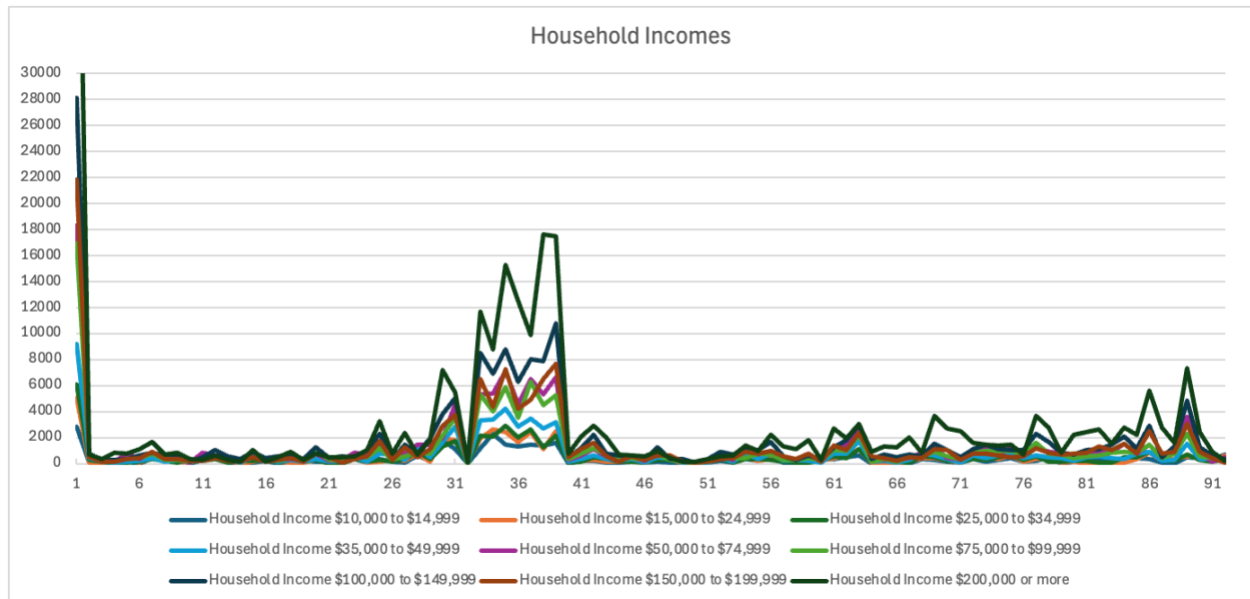
.

Vacant Housing Units

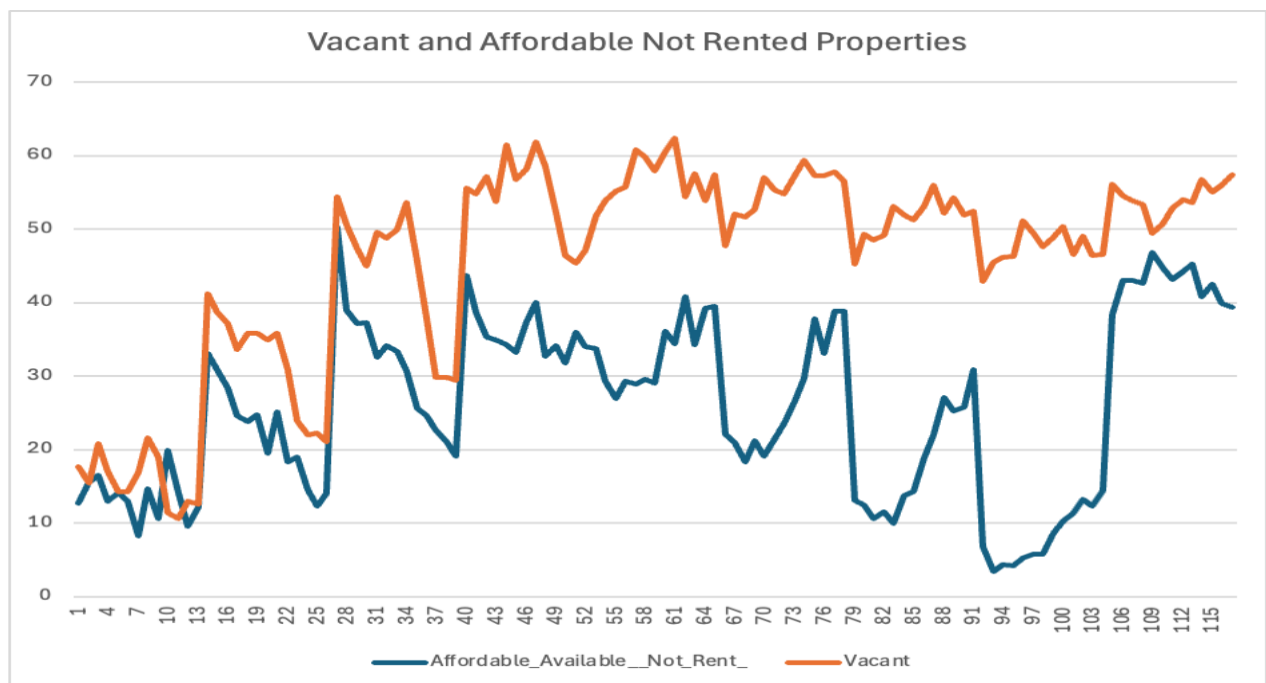
Percentiles		Smallest		
1%	7	7		
5%	50	41		
10%	92	41	Obs	92
25%	181	45	Sum of wgt.	92
50%	349.5		Mean	886.6304
		Largest	Std. dev.	1465.561
75%	826	4018		
90%	2530	4816	Variance	2147868
95%	3542	6632	Skewness	3.36136
99%	9359	9359	Kurtosis	16.36766



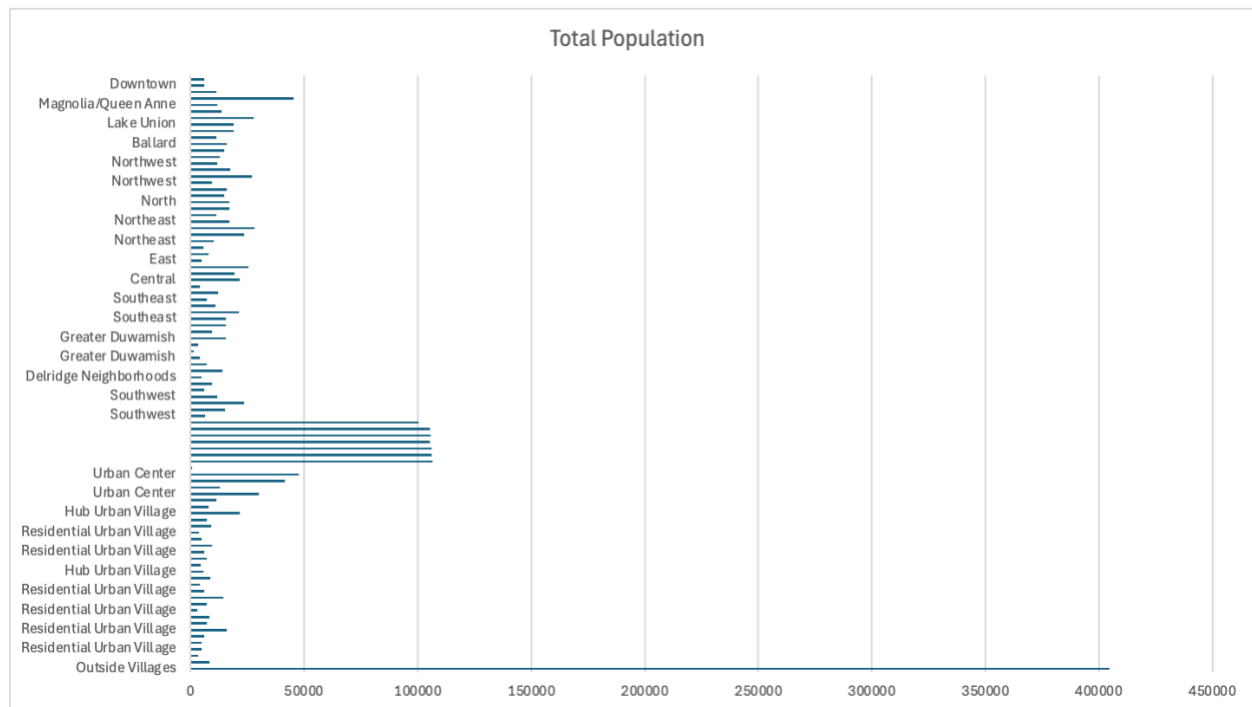
Occupations is an independent variable, because of the amount of income generated. Most income data points fall within a certain range in each category. Therefore, occupation is not interdependent on where a neighborhood is since the occupation pays around the same in each category.



Income is an independent variable, this is because across all neighborhoods, there is a mix of incomes. If income was interdependent, then the number of incomes would be more linear and not have data points that are spread apart.



One possible issue is that there are a lot of vacant properties, but they also share the same points as affordable but are not rented properties. Vacant and affordable are in separate columns but can report similar results.



Populations in neighborhoods can affect data outcomes, neighborhoods with high populations can have a wider range of incomes. Other neighborhoods with smaller populations could have a smaller range of incomes.