Chris Querfurth 000729686

Income Inequality In Chicago

Introduction

Income Inequality in the United States is often closely linked with demographic patterns,family dynamics and other broad socioeconomic factors. The Glass Ceiling[1] is a known phenomenon where women are underrepresented in higher income levels. Chicago is a Major US city known for its distinct neighborhoods with distinct demographic makeups (ethnic groups, poverty concentration, etc)

In this project, I wanted to analyze whether the gender composition of each neighborhood in Chicago is associated with its household income distribution.

I was interested in this question because I am a lifelong resident of Chicago, and have lived in both very poor neighborhoods and very affluent ones. I chose to look at gender's relationship to neighborhood income because I think that the relationship between poverty and race has been well documented and understood by the general public. Gender and Income is often discussed in terms of individual income but the extent to which neighborhood level gender composition correlates with income concentration is not frequently discussed.

Dataset

I started this project by searching for Chicago-centric datasets on kaggle. And found this 2023 American Community Survey[2]. It contains demographic information for the 77 Chicago neighborhoods. It includes breakdowns by income bracket, age/gender, and ethnicity. It is important to note that entries in the dataset relate to household/family income, not individual income. Below is a sample of the dataset.

	ACS Year	Community Area	Under $25,000	$25,000 to $49,999	$50,000 to $74,999	$75,000 to $125,000	$125,000 +	Male 0 to 17	Male 18 to 24	Male 25 to 34	Male 35 to 49	Male 50 to 64	Male 65+	Female 0 to 17	Female 18 to 24	Female 25 to 34	Female 35 to 49	Female 50 to 64	Female 65 +	Total Population	White	Black or African American	American Indian or Alaska Native	Asian
0	2023	ALBANY PARK	1269	1916	1801	2306	3379	4799	2955	4513	5442	4354	2287	4913	2405	4116	5228	3764	3054	47830	21496	2228	759	7124
1	2023	ARCHER HEIGHTS	223	752	441	795	739	1927	732	1102	1240	1417	1035	1502	899	978	1167	854	1021	13875	6232	10	108	679
2	2023	ARMOUR SQUARE	701	798	370	637	597	1300	487	871	1174	1177	1032	1022	517	965	1163	1378	2063	13149	2556	1487	107	8402
3	2023	ASHBURN	797	1351	1985	3014	2735	5150	1964	2881	4178	4228	2473	5138	2248	2774	4582	4044	3184	42842	11297	18124	697	436
4	2023	AUBURN GRESHAM	2541	2451	1592	2202	1850	5803	1836	2964	3431	3469	3223	5305	2304	3057	4522	5213	5357	46483	760	43414	119	399

Preliminary Analysis

Taking a look at some of the patterns in the dataset showed a large variance in the various demographic categories. Surprisingly the largest income bracket was >$125k and the smallest income bracket was <$25k. This makes some more sense when looking at the racial breakdown of the respondents, the vast majority are White, next Black and Asian/Multiracial/Others make up a much smaller population of respondents.
Taking a look at Gender, we see that most neighborhoods have around 50% male and female populations. Females slightly outnumber males across the city, with the biggest gap being in the 65+ age bracket. This makes logical sense given that women tend to live longer than men. I calculated a "% Female" statistic for each neighborhood by combining the counts for each Female age bracket and dividing by the Total Population of that neighborhood. To ensure that comparisons were fair across neighborhoods with different population sizes I calculated the percentage distribution of each income bracket within each neighborhood as well.

---

[1] https://bfi.uchicago.edu/wp-content/uploads/WP_2018-38.pdf
[2] https://www.kaggle.com/datasets/aniket0712/acs-5-year-data-by-community-area

Chris Querfurth 000729686

Taking a look at the correlation heatmap between the percentage of female population and income brackets the heatmap shows a positive relationship between % Female and the households earning under $25,000 and a negative relationship between % Female and households earning over $125,000. The middle income brackets show weak or near zero relationships.

Regression Analysis

I ended up running five OLS regression analyses. In each regression the Independent variable was the % Female and the Dependent variable was the percentage of households in a specific income bracket. For each regression the H_0: $\beta_1 = 0$ and H_A: $\beta_1 \neq 0$ Meaning no relationship between % Female and Income vs Percent Female does have an impact on Income. Results of all five regressions are below.

Income Bracket	Slope (β_1)	p-value	R^2	Decision ($\alpha = 0.05$)
Under $25,000	+1.4322	2.77e-06	0.255	**Reject H_0**
$25,000 to $49,999	+0.3343	0.1624	0.026	Fail to Reject H_0
$50,000 to $74,999	–0.0171	0.9171	0.000	Fail to Reject H_0
$75,000 to $125,000	–0.0937	0.5837	0.004	Fail to Reject H_0
Over $125,000	–1.6557	0.00638	0.095	**Reject H_0**

Key Findings

For the Under $25,000 analysis we reject the null hypothesis and confirm that there is a positive relationship between % Female in a neighborhood and amount of low-income households. For every percentage point that women exceed men in population, the amount of low-income households increases by 1.432%.

For Over $125,000, we reject the null hypothesis and confirm that there is a negative relationship between % Female and high income households. That is, for each percentage point that women exceed men in a neighborhood's population, the amount of households earning more than $125,000 decreases by 1.656%

We failed to reject the null hypothesis for the middle three income brackets. This means there was insufficient evidence to prove there was any relationship between a neighborhood's gender composition and household income.

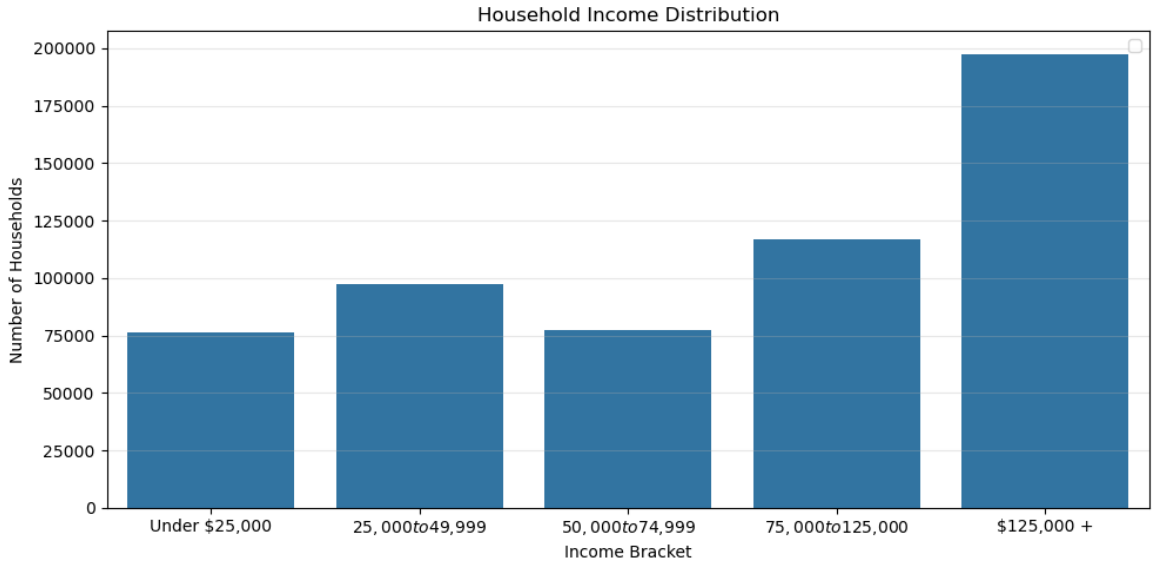
Conclusion

I was surprised by the results of the regression analyses at first. The idea that if a neighborhood has more women than men, they tend to both have fewer high income households and more low income households. However this result does support the idea of the Glass Ceiling. Women tend to have fewer opportunities for advancement and that on average they tend to make less money than men for the same work. I will note again that the dataset only shows household income but does not include size or structure of those households. It would be interesting to see how many households are dual-income or single, kids or no kids. It could stand to reason that married families generally have higher incomes and because there is both a man and a woman, the gender distribution tends to stay closer to 50%. At the same time I can also see how single women, or mother-only households earn less money comparatively.

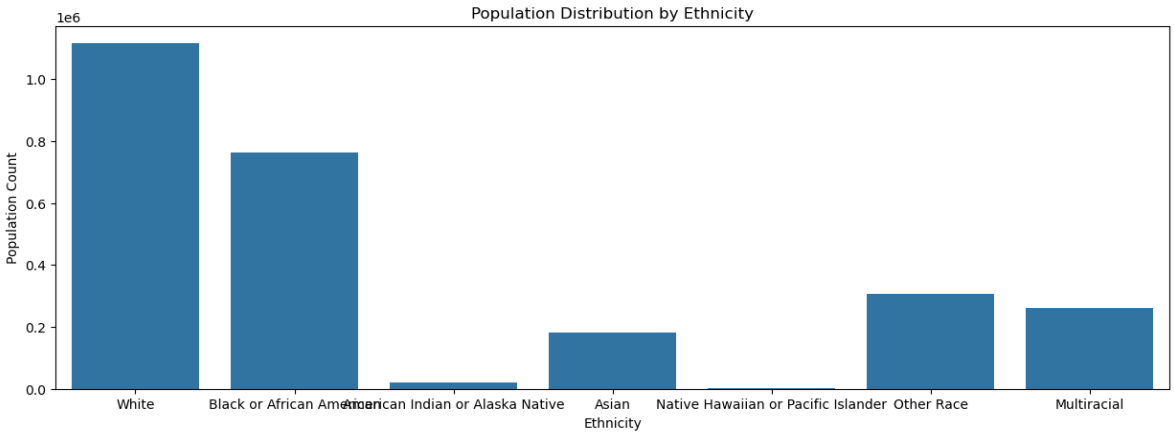
Chris Querfurth 000729686

Visualizations

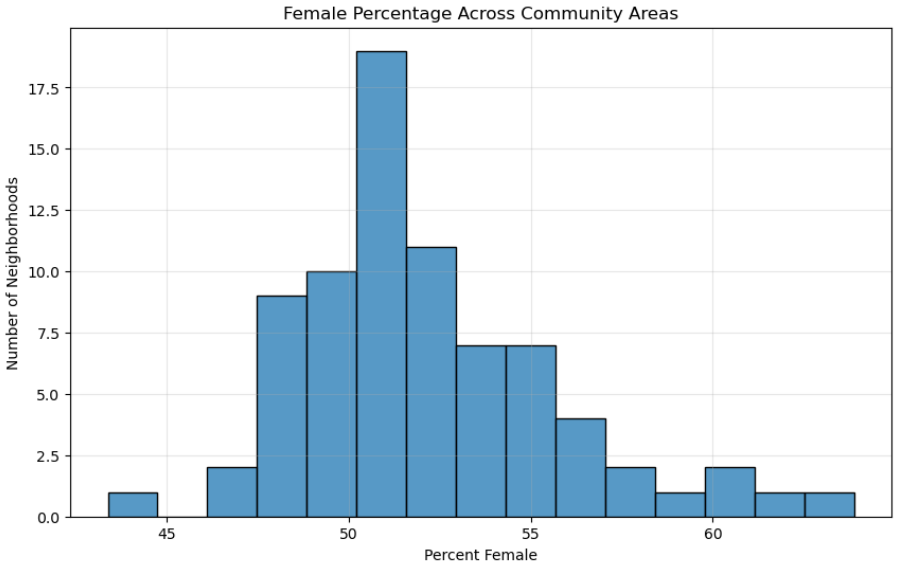
Income Distribution



Ethnicity Distribution

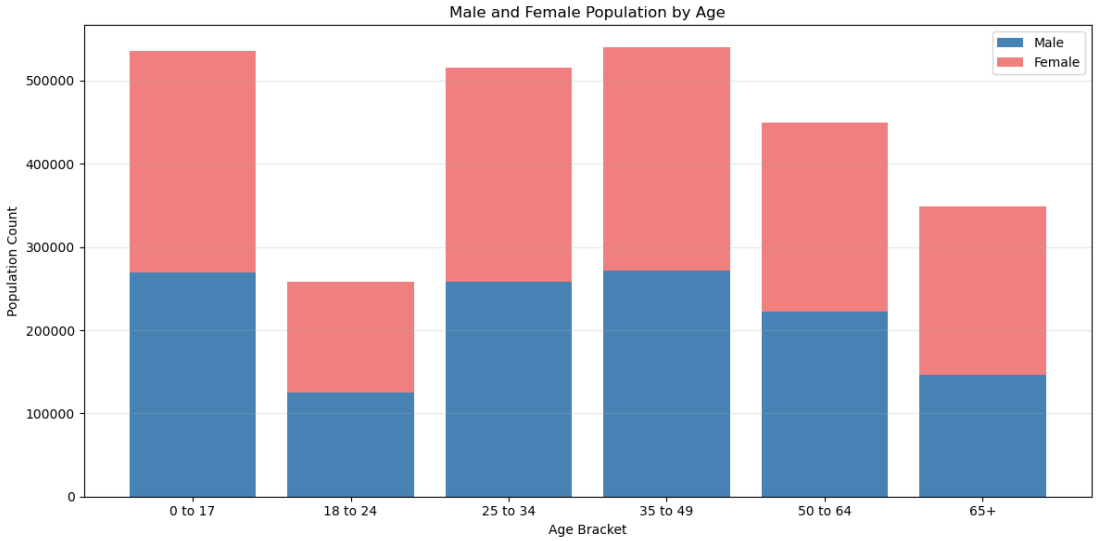


% Female Distribution

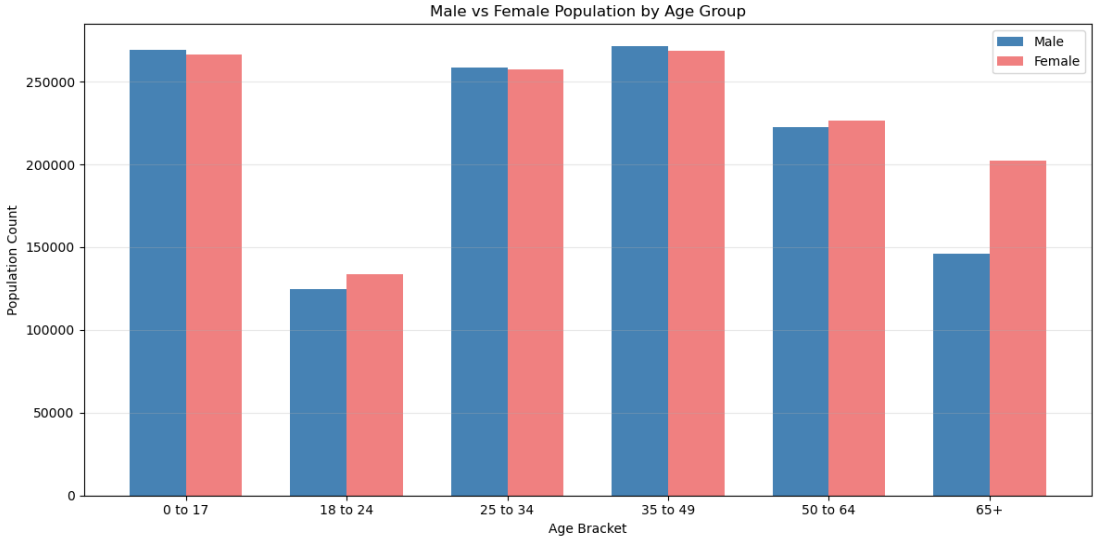


Chris Querfurth 000729686

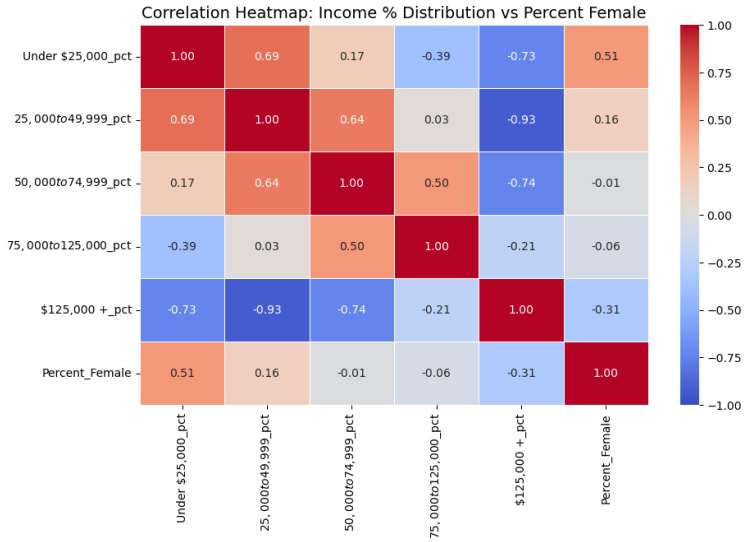
Male vs Female Population by Age Brackets



Male vs Female Population by Age Brackets



Correlation Heatmap



Chris Querfurth 000729686

Regression Analyses