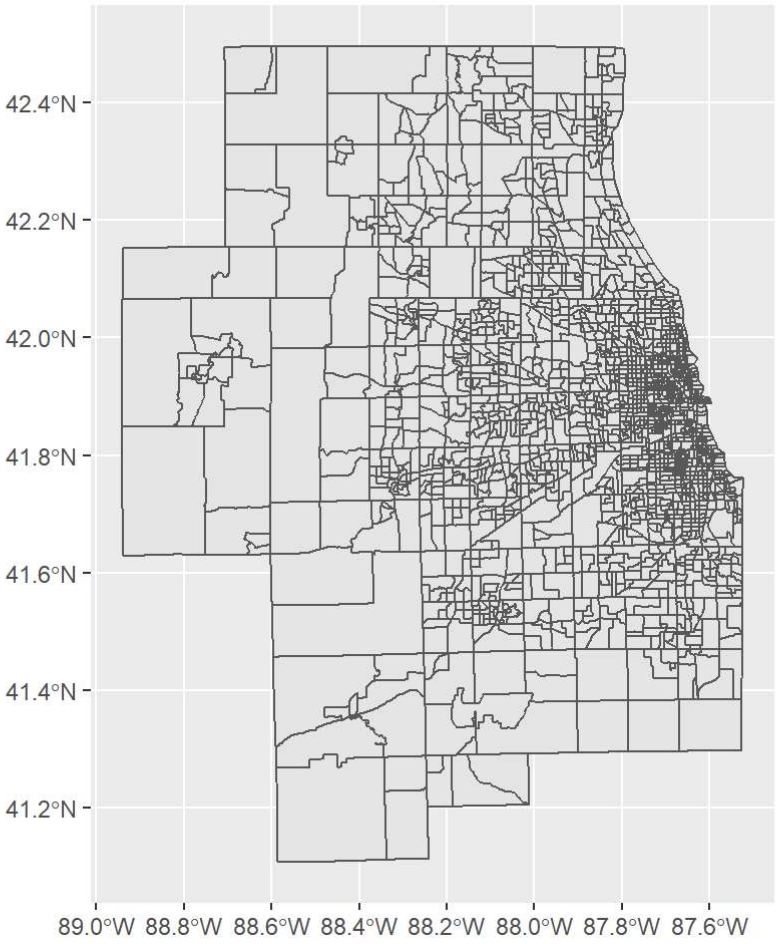
title: “Session 13 Polished Notebook” Author: “Rosario Barrera, DURP Student” Date: “March 16, 2021” output: html_document: df_print: paged pdf_document: default html_notebook: null —



Introduction

Segregation remains an enduring characteristic for most America cities because the root of the problem has not been solved. More specifically, segregation at its core continues to be about race, generational wealth, and capitalism. At its core, integration is nearly impossible if that means that families of specific races or ethnicity need to sacrifice specific social safety nets, savings, or their community in order to be part of an integrated community. Of course, these can be addressed through funding and support from local organizations and local governments. The wealth gap between black and white families continues to remain large which is an issue within itself. Despite efforts to reduce segregation, cities in America have remained that way because historical issues have not been adequately addressed at the education or at an institutional level. Segregation is not just about segregation, its also about the racial wealth gap, the education gap, generational wealth, lack of housing, and lack of resources (to name a few). The session 13 lab focuses on analyzing segregation throughout the Chicago land area. Segregation was measured using the dissimilarity index, the isolation index, interaction index,and how that correlates with income.The following will show the different ways in which the Chicago land area is segregated.

Racial Distribution

```
ggplot()+
  geom_sf(data=chi, aes(fill = PWhite, colour = PWhite))+
  geom_sf(data=co, fill=NA)+
  theme_minimal()+
  labs(title = " Chicago Land Area Percent White Distribution", subtitle = "Source: U.S Census 5
- Year ACS
       Figure 1")
```

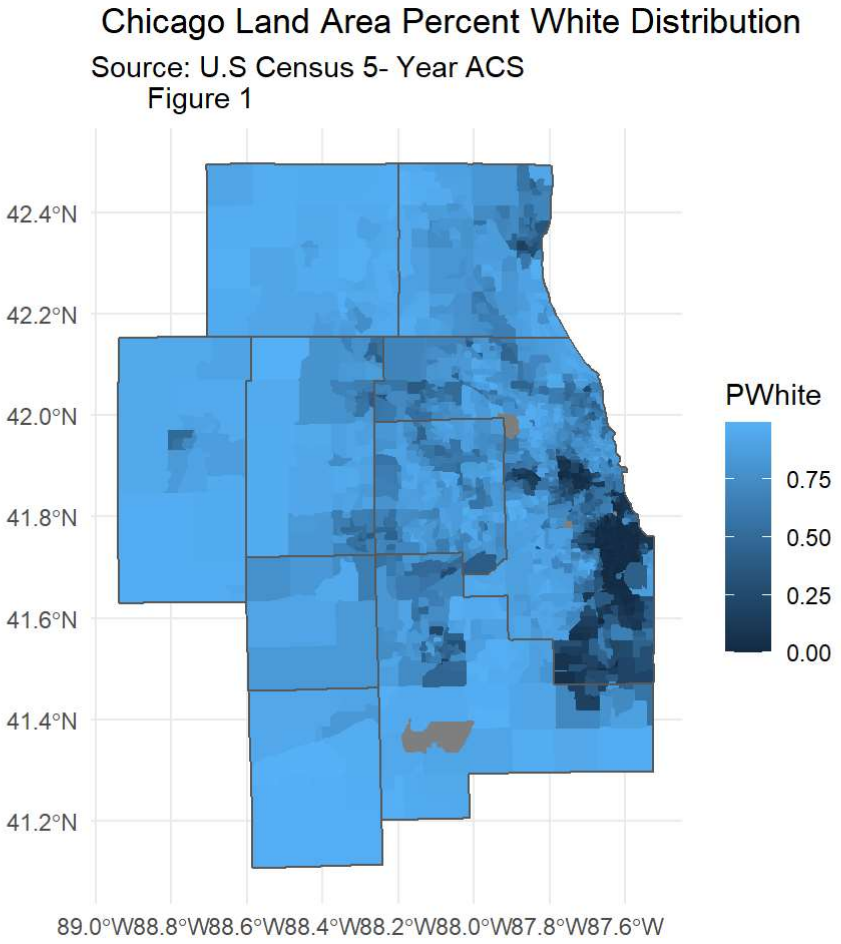
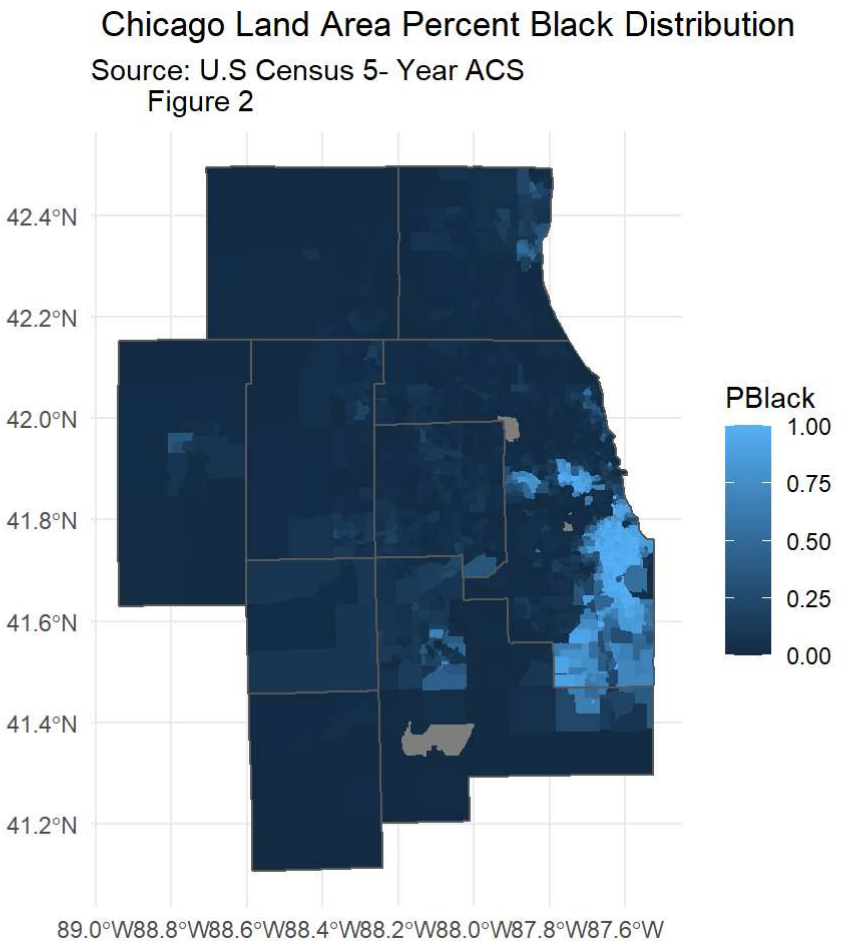
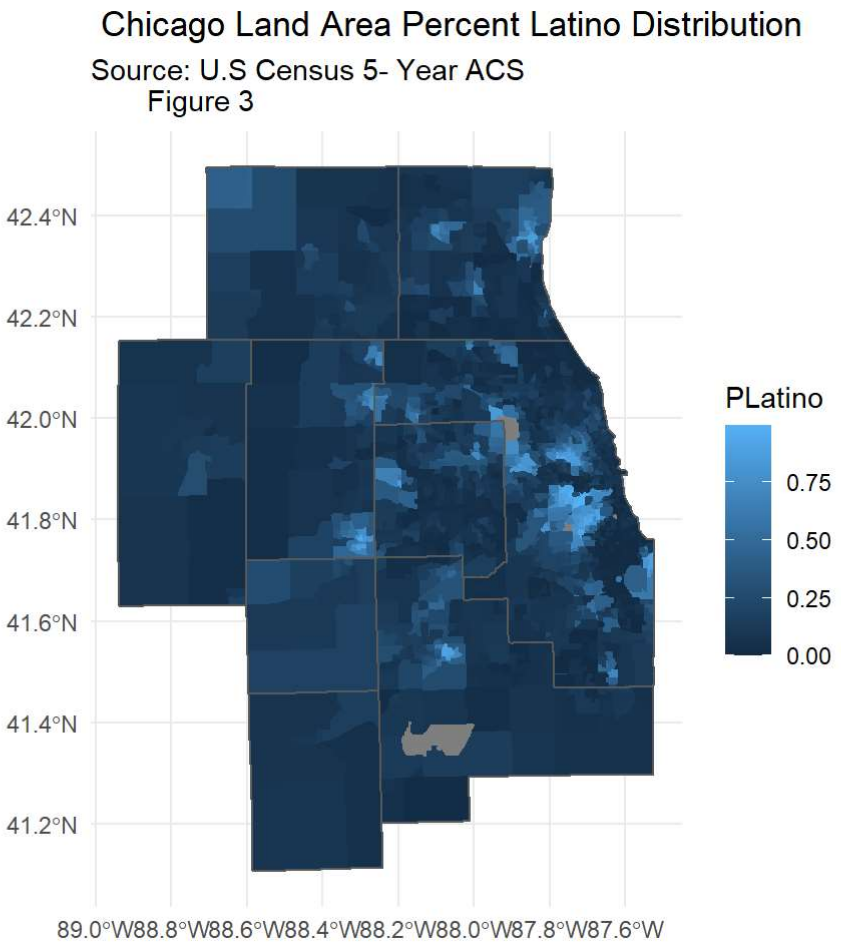


Figure 1 visualizes the distribution of White residents throughout the Chicago Land Area. One will see that the majority of these Counties have a high percentage rate of white residents. However, there area some areas the have a low percentage rate of white residents. Using the same mapping methods one can find out how many residents belong to to other racial groups throughout this region.

```
ggplot()+
  geom_sf(data=chi, aes(fill = PBlack, colour = PBlack))+
  geom_sf(data=co, fill=NA)+
  theme_minimal()+
  labs(title = " Chicago Land Area Percent Black Distribution", subtitle = "Source: U.S Census 5
- Year ACS
       Figure 2")
```




```
ggplot()+
  geom_sf(data=chi, aes(fill = PLatino, colour = PLatino))+
  geom_sf(data=co, fill=NA)+
  theme_minimal()+
  labs(title = " Chicago Land Area Percent Latino Distribution", subtitle = "Source: U.S Census
 5- Year ACS
       Figure 3")
```




```
ggplot()+
  geom_sf(data=chi, aes(fill = PAsian, colour = PAsian))+
  geom_sf(data=co, fill=NA)+
  theme_minimal()+
  labs(title = " Chicago Land Area Percent Asian Distribution", subtitle = "Source: U.S Census 5
- Year ACS
       Figure 4")
```

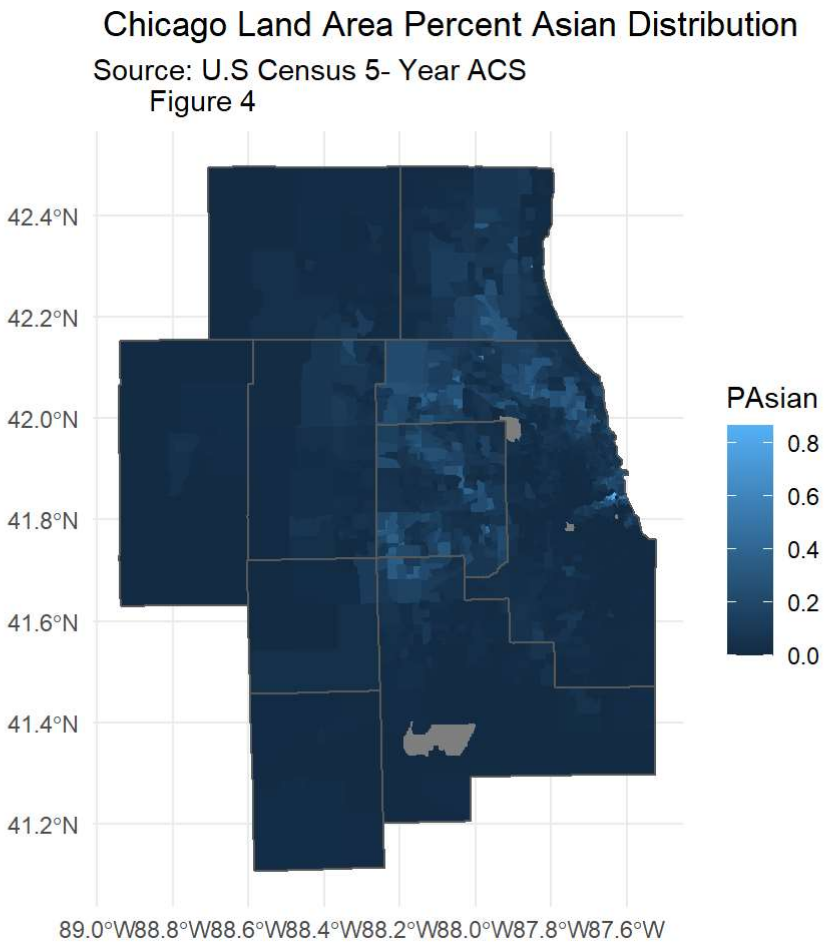


Figure 2 shows the percent distribution of the Black population in the Chicago Land Area. Figure 3 Shows the percent distribution of the Latino Population in the Chicago Land Area. Figure 4 shows the percent distribution of the Asian population. It is evident that within Chicago there are larger clusters of Latino and Black communities that don't quite exists in other Counties. One will also see that the Latino population is generally more dispersed throughout the Counties. For the Asian community, it appears that they make a small portion of the population, however, they are geographically located in particular areas in the region. For the Black population there is one general cluster to the south of the city . In general there is a lack of Black clusters throughout this region overall.

Measuring Racial Segregation

Maps 1, 2, 3 and 4 provide a general idea of where different racial and ethnic groups are located throughout the Chicago land area, however, they do not calculate how segregated these areas are. Using Census tracts as a unit of analysis, we will use this data to calculate segregation using the dissimilarity index,isolation index, and income as a factor to measure segregation.

Dissimilarity Index

" *Dissimilarity* is a common measure of *evenness* between two populations - dissimilarity measures the distribution of a minority population within a majority population. Conceptually, dissimilarity measures the proportion of that minority or subgroup population that would need to move in order to be equally distributed with the majority population. Dissimilarity is calculated as follows:

$D = .5 * \sum_i |\frac{b_i}{B} - \frac{w_i}{W}|$ where b_i is the number of blacks in tract *i* B is the number of blacks for the county w_i is the number of whites in tract *i* W is the number of whites for the county" (Greenlee Andrew, Session 13 Notebook, 2021)

Using the following definition I have calculated the dissimilarity Index between the White, Black, Asian, and Latino Population.A dissimilarity index number closer to one indicates a high level of segregation as the groups are extremely dissimilar.

Dissimilarity Between Black and White Residents

The following summary table shows the dissimilarity score between Black and White residents.

```
chi %>% mutate(dissim_wb = abs(Black / Black_co - White / White_co)) %>% group_by(COUNTYFP) %>%
summarise(dissimilarity = .5*sum(dissim_wb))
```

COUNTYFP <chr>	dissimilarity <dbl>	geometry <s_POLYGO>
031	0.7578794	<s_POLYGO>
037	0.5295320	<s_POLYGO>
043	0.3879989	<s_POLYGO>
063	0.3819132	<s_POLYGO>
089	0.4353399	<s_POLYGO>
093	0.2309396	<s_POLYGO>
097	0.5642134	<s_POLYGO>
111	0.3923875	<s_POLYGO>
197	0.5182092	<s_POLYGO>

9 rows

One will see that there is a wide distribution of dissimilarity between the counties. The County with the highest dissimilarity score is 031 which is cook county, and the one with the lowest is 093 which is Kendall County. This means that in Cook County (031) 75 percent of the Black population would have to relocate in order to achieve even distribution. This aligns with the visual in figure 2 as Black residents were concentrated within two evident clusters within the counties which was the southside of Chicago and further South West.

Dissmilarity Between Latino and White Residents

The following summary table shows the dissimilarity score between Latino and White Residents.

```
chi %>% mutate(dissim_lw = abs(Latino / Latino_co - White / White_co)) %>% group_by(COUNTYFP) %
>% summarise(dissimilarity = .5*sum(dissim_lw))
```

COUNTYFP <chr>	dissimilarity <dbl>	geometry <s_POLYGO>
031	0.4411871	<s_POLYGO>
037	0.2048503	<s_POLYGO>
043	0.3867571	<s_POLYGO>
063	0.2215250	<s_POLYGO>
089	0.4617327	<s_POLYGO>
093	0.1344544	<s_POLYGO>
097	0.4624586	<s_POLYGO>
111	0.2092301	<s_POLYGO>
197	0.4158243	<s_POLYGO>

9 rows

The dissimilarity score between White and Latinos is a lot less significant than it is between white and Black residents. This means that Latinos are more evenly distributed within these counties. This Confirms the visual data that was presented in Figure 3 as the Latino Population wasnt concentrated in one specific area.

Asian and White Dissimilarity

The following table demonstrates the dissimilarity between White and Asian residents

```
chi %>% mutate(dissim_lw = abs(Asian / Asian_co - White / White_co)) %>% group_by(COUNTYFP) %>%
summarise(dissimilarity = .5*sum(dissim_lw))
```

COUNTYFP <chr>	dissimilarity <dbl>	geometry <s_POLYGO>
031	0.4502416	<s_POLYGO>
037	0.4788275	<s_POLYGO>
043	0.3585515	<s_POLYGO>
063	0.4256758	<s_POLYGO>
089	0.3573122	<s_POLYGO>
093	0.4682907	<s_POLYGO>
097	0.3967334	<s_POLYGO>
111	0.4318553	<s_POLYGO>
197	0.5481525	<s_POLYGO>

9 rows

The dissimilarity between Asian and White residents appears to be even across all counties.This makes sense because figure 4 showed that Asian residents are quite evenly distributed across all counties with no apparent areas of concentration.

Interaction Index

"A second common measure of segregation is *interaction* which is a measure of *exposure*. Interaction measures the likelihood of population subgroups interacting with one another based upon their distribution within areal sub units (tracts). For instance, Black-White interaction is calculated as follows:

$Interaction = \sum_i \frac{b_i}{B} * \frac{w_i}{t_i}$ Where b_i is the Black population of tract *i* B is the Black population of the county w_i is the White population of tract *i* t_i is the total population of tract *i*" (Greenlee, Andrew Session 13 Notebook, 2021)

Black and White Interaction

The following summary table demonstrates the interaction score between White and Black residents

```
chi %>% mutate(interaction_bw = (Black / Black_co * White / Pop)) %>% group_by(COUNTYFP) %>% sum
marise(interaction = sum(interaction_bw, na.rm=TRUE))
```

COUNTYFP <chr>	interaction <dbl>	geometry <s_POLYGO>
031	0.2106465	<s_POLYGO>
037	0.6671092	<s_POLYGO>
043	0.6856959	<s_POLYGO>
063	0.9006916	<s_POLYGO>
089	0.6272239	<s_POLYGO>
093	0.7817260	<s_POLYGO>
097	0.5771504	<s_POLYGO>
111	0.8959265	<s_POLYGO>
197	0.5353900	<s_POLYGO>

9 rows

One will see that the interaction index scores vary significantly. On the higher end in county 063, black and white residents are 90% likely to interact. However, on the lower end, in county 031, black and white residents are only 20% likely to interact. This is significant because considering that Chicago is located in this county, it is very telling about the segregation that exists within the City of Chicago.

White and Latino Interaction

The following summary table demonstrates the interaction score between White and Latino Residents.

```
chi %>% mutate(interaction_bw = (Latino / Latino_co * White / Pop)) %>% group_by(COUNTYFP) %>% s
ummarise(interaction = sum(interaction_bw, na.rm=TRUE))
```

COUNTYFP <chr>	interaction <dbl>	geometry <s_POLYGO>
031	0.5782081	<s_POLYGO>
037	0.8334113	<s_POLYGO>
043	0.7284163	<s_POLYGO>
063	0.9158459	<s_POLYGO>
089	0.6022007	<s_POLYGO>
093	0.8049897	<s_POLYGO>
097	0.6588083	<s_POLYGO>
111	0.9159635	<s_POLYGO>
197	0.6438372	<s_POLYGO>

9 rows

The interaction index scores between the White and Latino population are generally much higher than the scores between White and Black Residents. In general, the scores are above 55% likelihood of White and Latino residents interacting with one another. At this point of the segregation analysis it is evident that Black residents are most likely to experience the effects of segregation whereas Latinos appear to be more evenly distributed through the Chicago land Counties. In general, Latinos are very likely to interact with white residents.

White and Asian Interaction

The following summary table demonstrates the interaction score between WHite and Asian Residents

```
chi %>% mutate(interaction_bw = (Asian / Asian_co * White / Pop)) %>% group_by(COUNTYFP) %>% sum
marise(interaction = sum(interaction_bw, na.rm=TRUE))
```

COUNTYFP <chr>	interaction <dbl>	geometry <s_POLYGO>
031	0.6191678	<s_POLYGO>
037	0.7229024	<s_POLYGO>
043	0.6921109	<s_POLYGO>
063	0.9161974	<s_POLYGO>
089	0.7511793	<s_POLYGO>
093	0.7261192	<s_POLYGO>
097	0.7328624	<s_POLYGO>

COUNTYFP <chr>	interaction <dbl>	geometry <s_POLYGO>
111	0.8873723	<s_POLYGO>
197	0.6430860	<s_POLYGO>

9 rows

The interaction index scores between the White and Asian population are generally high. They are the highest of all ethnic/ racial groups analyzed. In general Asian residents have a 61% chance or higher of interacting with White residents.

Isolation Index

Another measure of segregation is "*isolation* which is only measured for one group at a time. Isolation measures the likelihood of contact for a subgroup with other subgroup members. It is interpreted similarly to Interaction (as a probability ranging from 0 to 1):

#Isolation for Black residents $Isolation = \sum_i \frac{b_i}{B} * \frac{b_i}{t_i}$ Where b_i is the Black population of tract *i* B is the Black population of the county t_i is the total population of tract *i*" (Greenlee, Andrew, Session 13 Notebook, 2021)

Black Isolation

The following summary table shows the isolation scores for Black residents in the Chicago MSA

```
chi %>% mutate(interaction_bw = (Black / Black_co * Black / Pop)) %>% group_by(COUNTYFP) %>% sum
marise(interaction = sum(interaction_bw, na.rm=TRUE))
```

COUNTYFP <chr>	interaction <dbl>	geometry <s_POLYGO>
031	0.69850248	<s_POLYGO>
037	0.21109286	<s_POLYGO>
043	0.09575918	<s_POLYGO>
063	0.02351011	<s_POLYGO>
089	0.10711290	<s_POLYGO>
093	0.08875501	<s_POLYGO>
097	0.20231640	<s_POLYGO>
111	0.03060264	<s_POLYGO>
197	0.28654933	<s_POLYGO>

9 rows

It is evident that Cook County (031) has a high isolation index for Black residents in the Chicago MSA. This aligns with what was analyzed earlier as Black residents experience the highest rates of dissimilarity. It is evident that Black residents remain significantly isolated from other other groups within the region.

Latino Isolation

The following summary table shows the isolation score of Latino Residents in the Chicago MSA

```
chi %>% mutate(interaction_bw = (Latino/ Latino_co * Latino / Pop))%>% group_by(COUNTYFP)%>%
summarize(interaction = sum(interaction_bw, na.rm = TRUE))
```

COUNTYFP <chr>	interaction <dbl>	geometry <s_POLYGO>
031	0.5318494	<s_POLYGO>
037	0.1337460	<s_POLYGO>
043	0.2842597	<s_POLYGO>
063	0.1205511	<s_POLYGO>
089	0.5424233	<s_POLYGO>
093	0.2040383	<s_POLYGO>
097	0.4278993	<s_POLYGO>
111	0.1789880	<s_POLYGO>
197	0.3042042	<s_POLYGO>

9 rows

The Latino isolation scores are generally lower. The higher isolation scores are in Cook County (031) and Kane County (089). As aforementioned, Latinos are less likely to experience segregation thank black residents in the Chicago MSA.

Asian Isolation

The following summary table shows the isolation score of Latino Residents in the Chicago MSA

```
chi %>% mutate(interaction_bw = (Asian/ Asian_co * Asian / Pop))%>% group_by(COUNTYFP)%>%
summarize(interaction = sum(interaction_bw, na.rm = TRUE))
```

COUNTYFP <chr>	interaction <dbl>	geometry <s_POLYGO>
031	0.20505869	<s_POLYGO>
037	0.04605683	<s_POLYGO>
043	0.18556734	<s_POLYGO>

COUNTYFP <chr>	interaction <dbl>	geometry <s_POLYGO>
063	0.01696605	<s_POLYGO>
089	0.07813193	<s_POLYGO>
093	0.08022315	<s_POLYGO>
097	0.15677809	<s_POLYGO>
111	0.05902123	<s_POLYGO>
197	0.16398013	<s_POLYGO>

9 rows

Of the three groups, the Asian population is the least isolated. This is on par with figure 4 as there was significant visual evidence of Asian residents clustering in one particular area of the region.

Income Inequality and Segregation

Another way people can be segregated is through income. A family or individuals income can determine where a person will live due to housing affordability. We can measure these disparities through measures the income inequality of the basis of median household income data.

```
#Loading data
B19013<-get_acs(geography = "county", state = state, variables = B19013_Vars, survey = survey, y
ear = DL_Year, output = "wide")
B19013$MHHI_co<-B19013$B19013_001E
B19013$MHHI_co[B19013$MHHI_co == "NaN"]<-NA
B19013<-B19013 %>% select(GEOID, MHHI_co)
chi<-chi %>% mutate(COUNTY = paste0(STATEFP, COUNTYFP))
chi <- left_join(chi, B19013, by =c ("COUNTY"= "GEOID"))
chi<-chi %>% mutate(inc_ratio = MHHI / MHHI_co)
```

```
ggplot()+
  geom_sf(data=chi, aes(fill=inc_ratio), colour = NA)+
    labs(title = "Income Ratio for the Chicago Region, Figure 5", fill = "Income Ratio")+
  geom_sf(data=co, fill=NA) +
  theme_minimal()
```

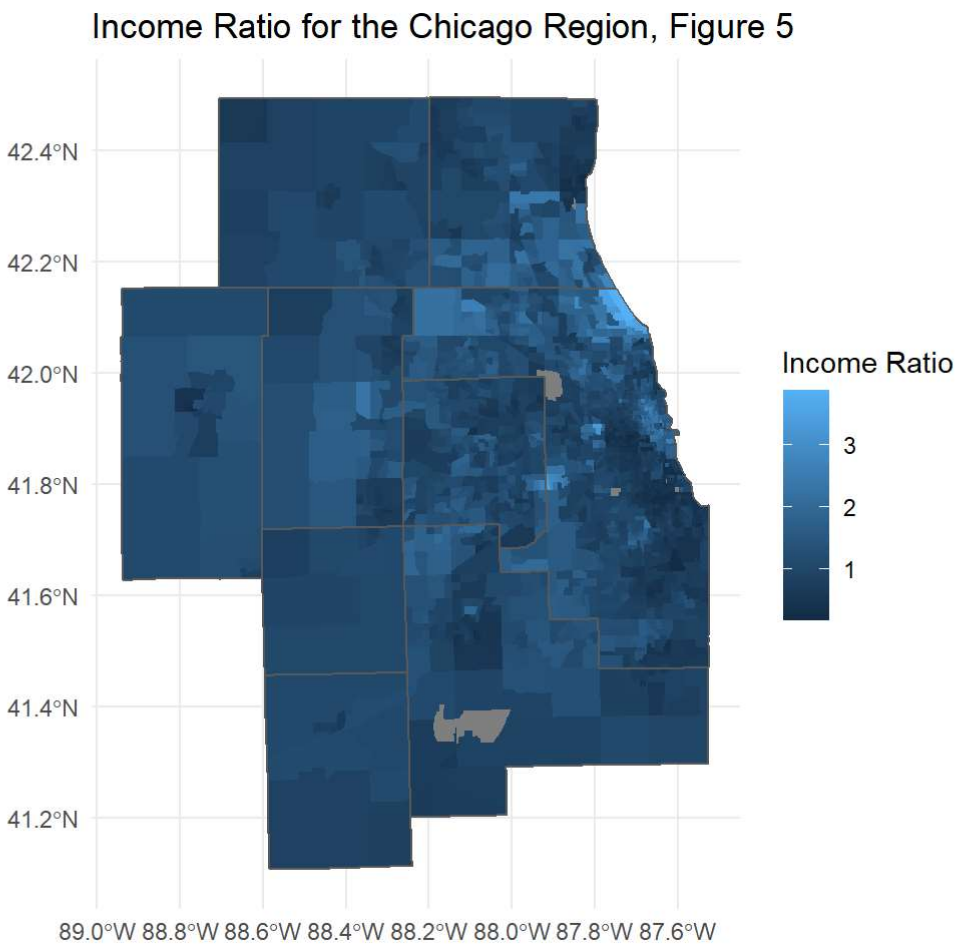
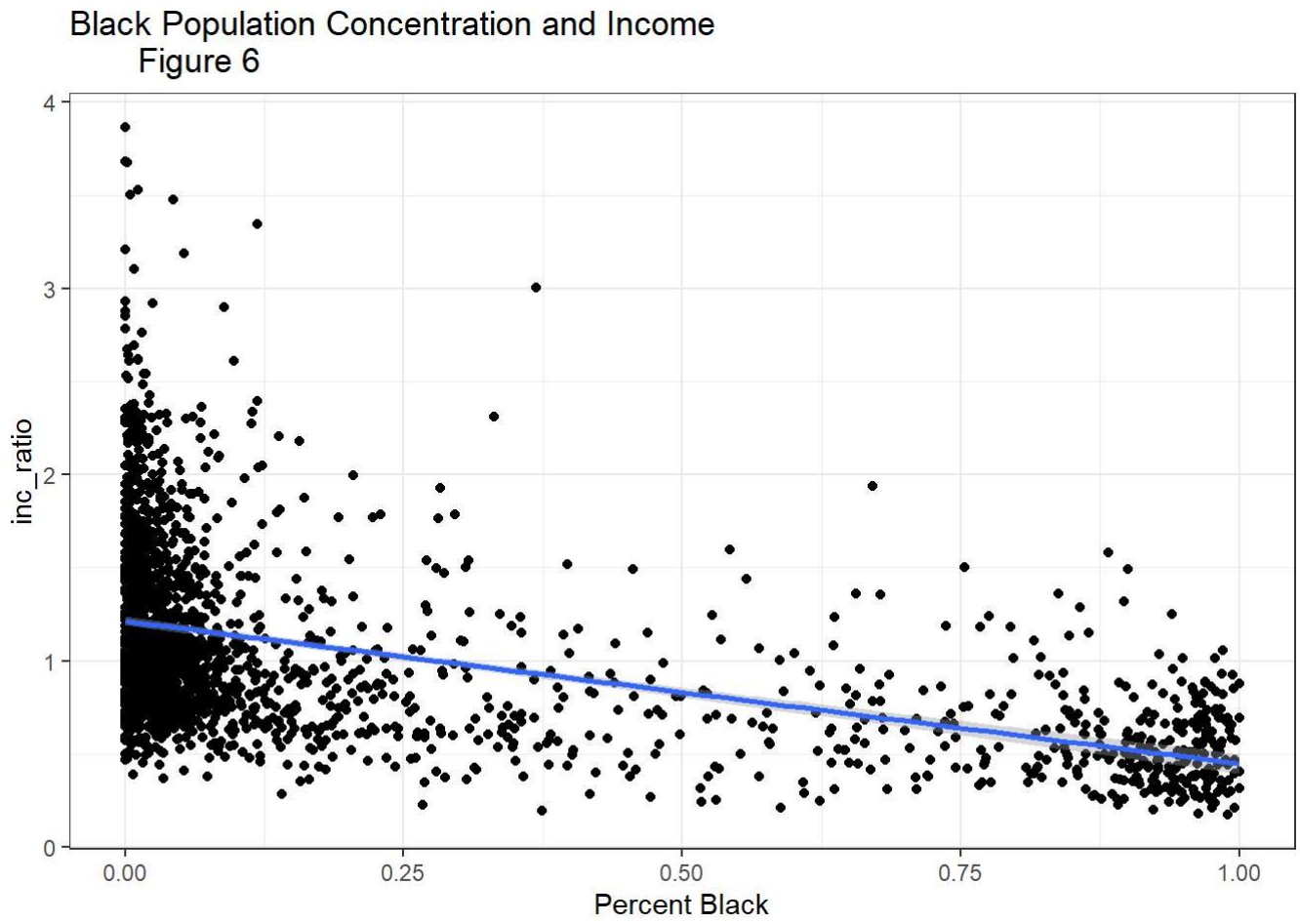


Figure 4 demonstrates the distribution of income throughout the Chicago MSA. It is evident that there are clusters in the southern region that live in the lower income distribution. When comparing this to figures 1 and 2 one will see that the lower income ratios are located where black residents are most concentrated, and white residents live in the higher income ratios. it is evident that throughout the Chicago MSA there is segregation based on race and income.

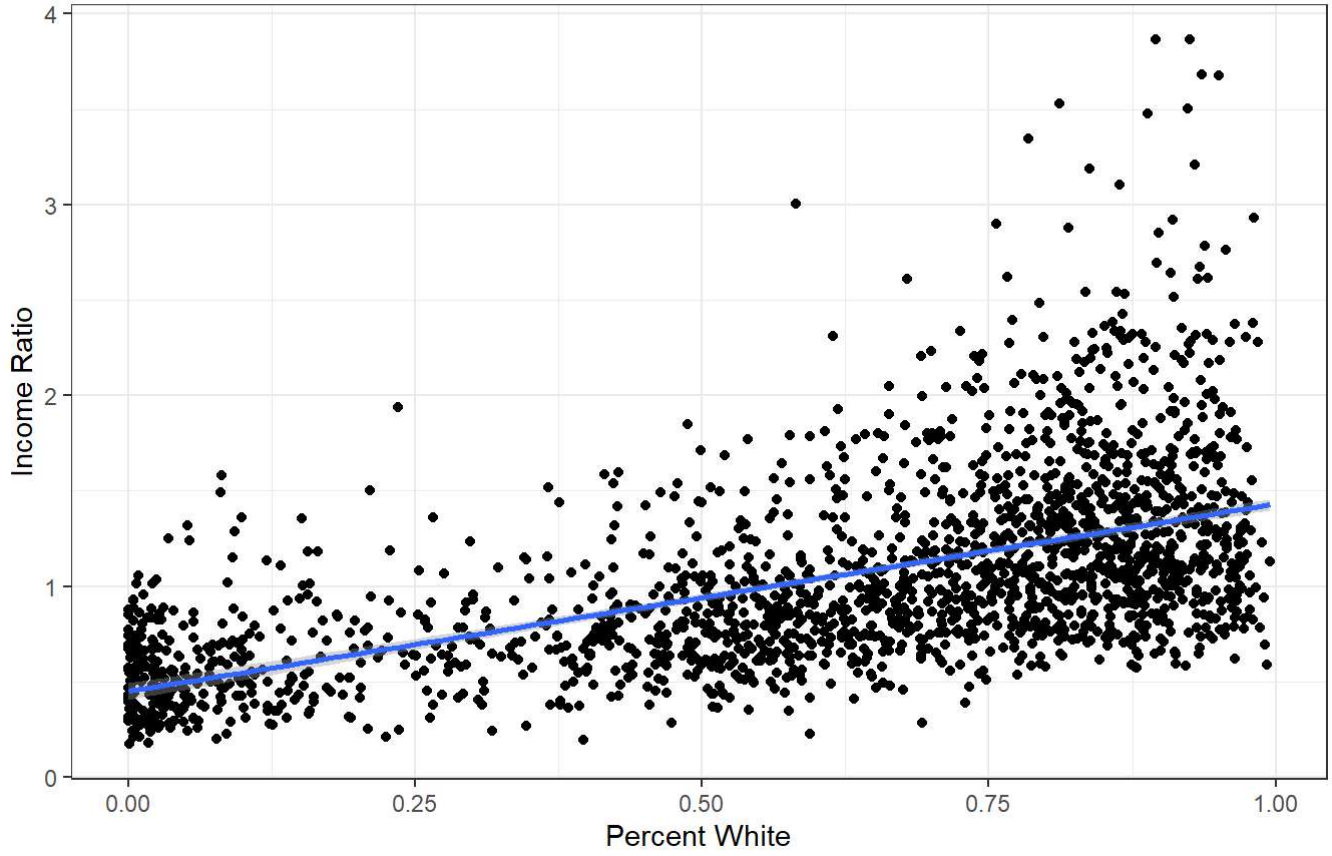
The following graphs demonstrate the relationship between racial concentration and income.

```
ggplot()+
  geom_point(data=chi, aes(x=PBlack, y=inc_ratio))+
  geom_smooth(data=chi, method = lm, aes(x=PBlack, y = inc_ratio))+
  theme_bw()+
  labs(title = "Black Population Concentration and Income
       Figure 6", x= "Percent Black", Y = "Income Ratio")
```

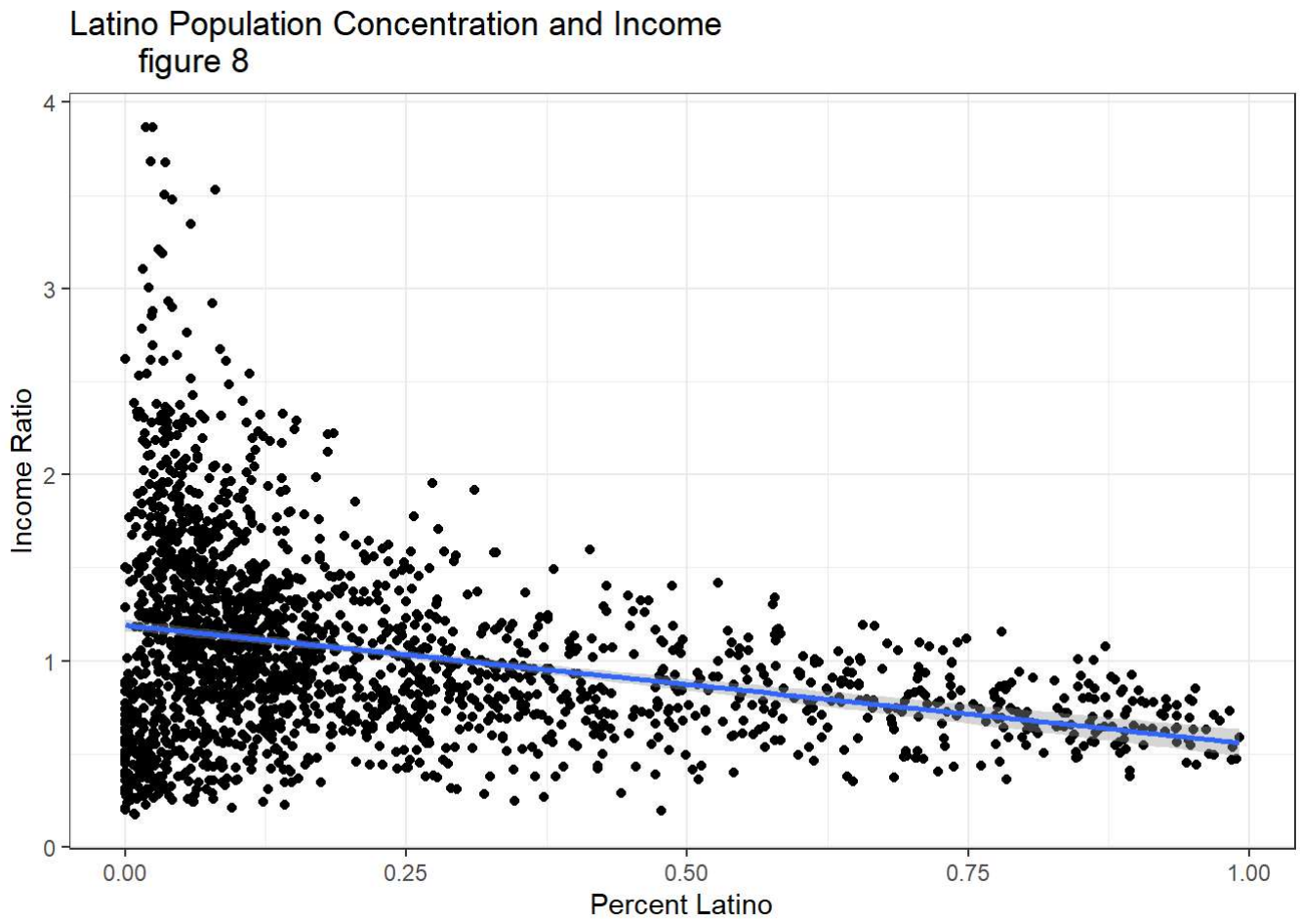



```
ggplot()+
  geom_point(data=chi, aes(x=PWhite, y=inc_ratio))+
  geom_smooth(data=chi, method = lm, aes(x=PWhite, y = inc_ratio))+
  theme_bw()+
  labs(title= "White Population Concentration and Income
       Figure 7", x = "Percent White", y = "Income Ratio")
```

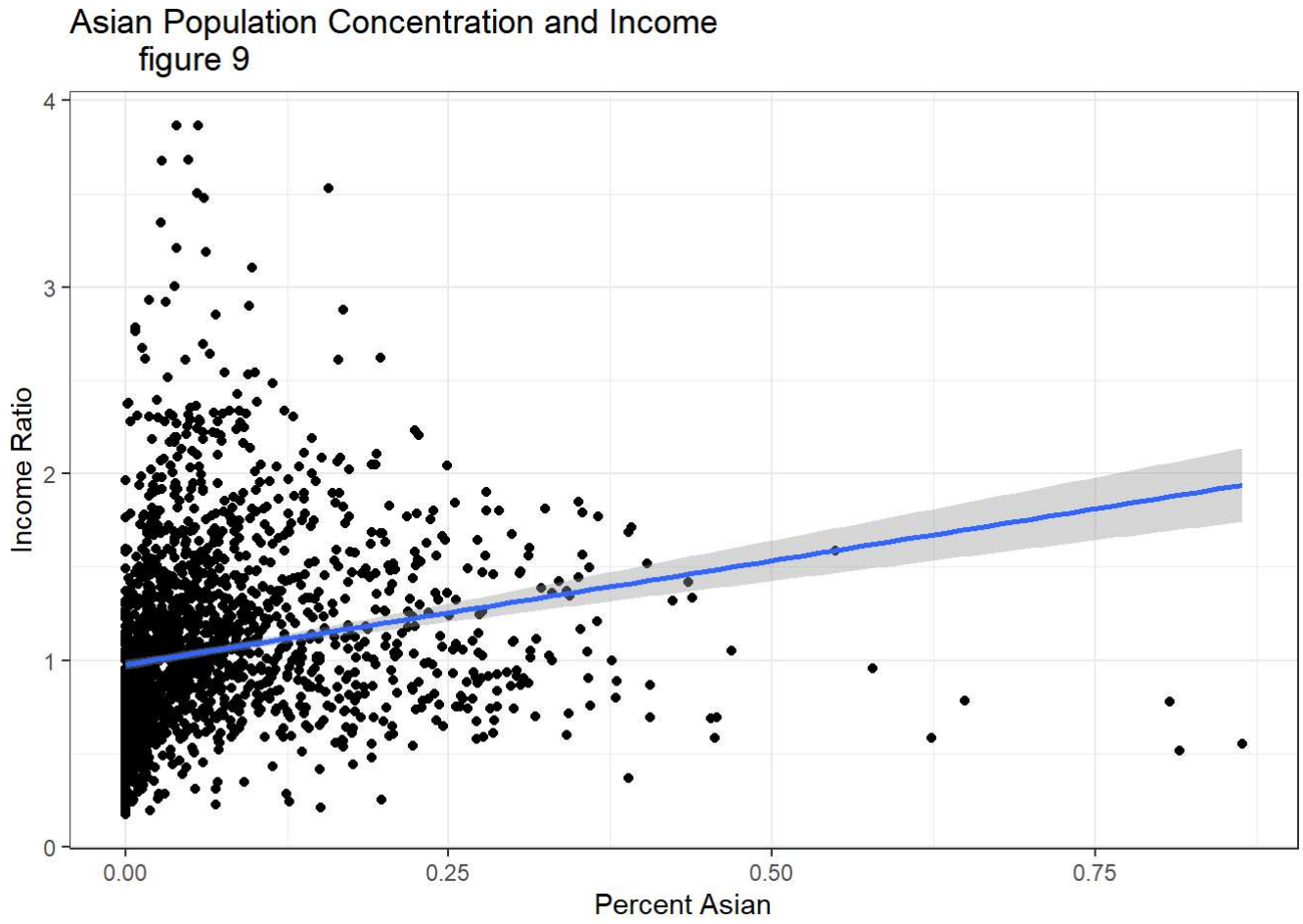
White Population Concentration and Income
Figure 7




```
ggplot()+
  geom_point(data=chi, aes(x=PLatino, y=inc_ratio))+
  geom_smooth(data=chi, method = lm, aes(x=PLatino, y = inc_ratio))+
  theme_bw()+
  labs(title= "Latino Population Concentration and Income
       figure 8", x = "Percent Latino", y = "Income Ratio")
```




```
ggplot()+
  geom_point(data=chi, aes(x=PAsian, y=inc_ratio))+
  geom_smooth(data=chi, method = lm, aes(x=PAsian, y = inc_ratio))+
  theme_bw()+
  labs(title= "Asian Population Concentration and Income
       figure 9", x = "Percent Asian", y = "Income Ratio")
```



Figures 6, 7, 8, and 9 are graphs that demonstrate the relationship between income and population concentration. There is evidence that shows that higher incomes are within communities where the percentage of white residents are higher. Conversely, lower incomes communities are those where black and Latino residents reside. For the Asian community incomes are generally higher as well. For the Asian community it appears that the less concentration of Asian there is, the higher their incomes are. This supports the hypothesis that counties and communities are also segregated by income in addition to segregation by race and ethnicity.

Conclusion

Segregation can be measured in various forms. We measured segregation through 4 different sets of measures and all concluded that segregation is existent in the Chicago MSA. More specifically, the residents that experience the most effects of segregation are Black Residents. In general, Black residents are concentrated within one geographical area. This concentration is seen in figure 1 where there is an explicit cluster of black residents in the South side of Chicago. Similarly, Black residents experienced the highest scores in all three indexes: the isolation index, the dissimilarity index, and interaction index.

Additionally, it was also evident that segregation also exists on the basis of income. Figure 6,7,8, and 9 show the relationship between income and population concentration. For areas with a high percentage of white residents income was higher. For areas with a high percentage of Black or Latino residents the incomes were lower/ decreased. When comparing the income ratio map (figure 5), we can also see the spatial distribution of incomes showing that lower income populations were concentrated in specific geographic areas.

While these segregation measures showed us that segregation patterns continue to exist there are other ways we can also explore segregation. People can be segregated by other factors such as educational attainment, housing type (single family, duplex, condo) based on housing affordability, social service uptake, and age to name a few.

However, the measures used in this lab are helpful in showing the existing conditions in terms of segregation which are useful for informing policy.