# A Close Look At Cincinnati and Covington

36315 Final Project Report

Aashna Singh. Mei Kuo. Nolan Carroll. Wan Xin Teo. Yijie Qiu.

## INTRODUCTION

Our cities of interest at the block-group level are Cincinnati and Covington. Cincinnati is a city and the county seat of Hamilton County in Ohio State that has a population of 296,943 according to the 2010 US Census, making it the third-largest city in Ohio. On the other hand, Covington is a city in Kenton County in Kentucky State that has a city population of 40,640 as of the 2010 US Census, making it the fifth-most-populous city in Kentucky. Since Covington and Cincinnati locate right beside each other geographically, we decide to combine the spatial polygons of Covington and Cincinnati together to give clearer and easier comparisons between two cities. Our goal of the report is to examine the distribution of population characteristics, including population, race, income and age distributions, across Cincinnati and Covington.

## DATA AND METHODS

Our data on Cincinnati and Covington come from the 2000 and 2010 United States Census, and the 2011 American Community Survey. Specifically, we explore the following questions with the corresponding graphs to visualize the difference:

- How population distributions of the four major ethnic groups differ in 2000 and 2010 for Cincinnati and Covington. The four major ethnic groups are Asian, African Americans, Hispanic and Caucasian. We show the change in population by race using barplots.
- How race distributions of African American and Caucasian changes in Cincinnati and Covington from 2000 to 2010. We use choropleth maps colored by the percentage change in race distributions.
- How income distributions of the four major ethnic groups correlate with population distributions in 2010 for Cincinnati and Covington. We use scatter plots to show the tract-wide median income by percentage of ethnic population with reference lines added for city-wide median income and average percentage of four major ethnic groups. In addition, we compare the income distributions in Cincinnati and Covington using choropleth map colored by median income in 2010.
- How age distributions differ for female and male in Cincinnati and Covington in 2010. We use the population pyramids to get the information about population-age structure by gender in Cincinnati and Covington.

## ANALYSIS

We begin our analysis on the population distributions by race in Cincinnati and Covington. All the corresponding graphs are attached in Appendix. In each section, we discuss why we have chosen the graph and not others as well as the features in each graph.

### Race Distribution Percentage for African American and Caucasian in Cincinnati and Covington between 2000 and 2010

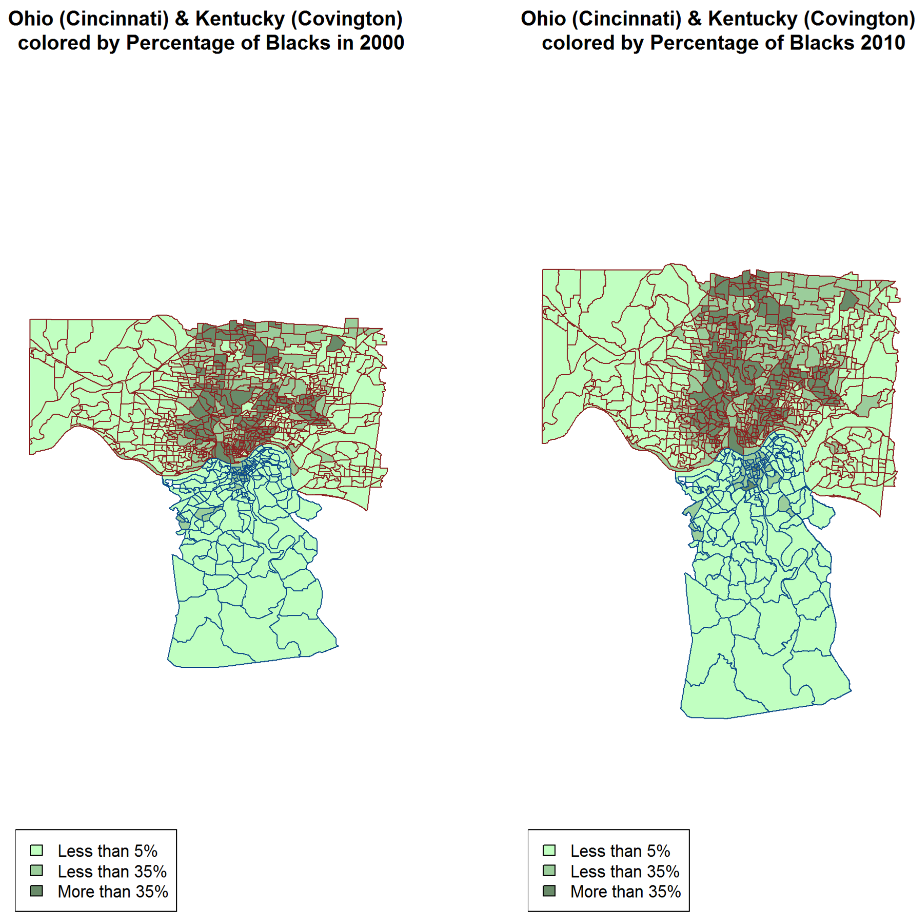



Figure 1 Choropleth of Cincinnati and Covington by percentage of African American population

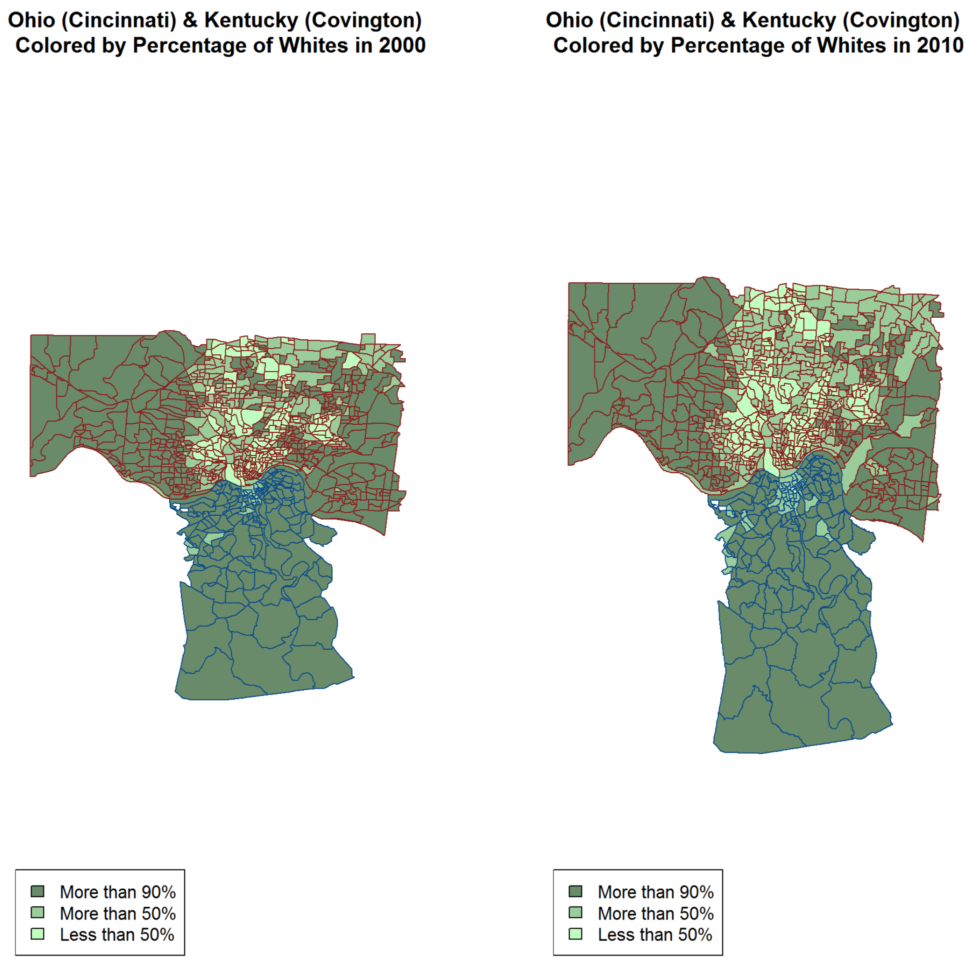



Figure 2 Choropleth of Cincinnati and Covington by percentage of Caucasian population

We plotted choropleth plots to show the distribution of race in 2000 and 2010. We focused on two main races- African American and Caucasian. We chose to plot these features on a map because it would be easier to compare cross border differences in race as well as within border differences in race. We also chose to use a gradient color scheme versus coloring based on bins. We did this because we thought it would be difficult for the user to keep track of which color meant which bin whereas if we used a gradient it would be easier to keep track of what a color indicated relative to the other colors. To decide the bins for the gradient we rounded up the 1st and the 3rd quartile as the cutoff points. In our gradient color scheme the darker the color got the higher number of observations in that area.

Because the tracts are so small we do not actually get to see the which tract had an increase/decrease in population of a race. In other words the choropleths from 2000 and 2010 do indicate change in distribution of race but the details are hard to read. To overcome this we could have plotted the change in every tract on a choropleth (such that it will be on one instead of 2 plots) and that would have been easier to understand. We could have plotted a density plot along the longitude (since we know that the African Americans seem to be concentrated in the middle of the plot) to see where most of the African Americans/Caucasians are concentrated. If there was a shift in distribution of race we would be able to see clearly how it changed.

In graph 1, the red border represents the Cincinnati tracts and the blue border represents the Covington tracts. In general, looking at both plots, we can see that Covington in general has less than 5% African Americans for most tracts whereas for Cincinnati, there are many tracts in the middle where there is a larger presence of African American population. There is an increase in the number of tracts with more than just 5% of population being African American for both cities, but it is more apparent in Cincinnati. These tracts that saw an increase in percentage of African American still tend to be in the middle of Cincinnati, but the area is greater (spreading to the left and right of the original area in 2000 where there was a larger presence of African American community). However, this spread in African American community is larger to the right of the plot compared to the left as we can see more tracts with greater than just 5% African American to the right of the plot. In Covington, the tracts that saw an increase in African American population to more than 5% tend to be closer to the border shared with Cincinnati and it is in the center of the border.

In graph 2, the red border represents the Cincinnati tracts and the blue border represents the Covington tracts. In general, looking at both plots, we can see that Covington in general has more than 90% whites for most tracts whereas for Cincinnati, there are many tracts in the middle where there is below 90% whites. There is a decrease in the number of tracts with more than 90% of population being white for both cities, but it is more apparent in Cincinnati. The tracts in the top right hand corner have generally 50-90% of whites in 2010 compared to mostly having more than 90% whites in 2010. While Covington has a few tracts that saw percentage of population being white drop below 90%, it happened in the tracts nearer to center of Cincinnati (where there are more tracts with lower percentages of whites).

It is interesting to see how Cincinnati and Covington share borders but have almost the opposite racial distribution in terms of african americans and caucasians.

**Percentage change in population for the main races in Cincinnati and Covington between 2000 and 2010**

To show the percentage change in race we decided to plot the change on a barplot. In a barplot we can compare across different races about how much one increased and decreased. In this case we get a number for % change and it is easy to visualize and understand the plot. One can easily look at the heights of the bars and compare which race had a greater change in population and whether it was negative or positive. However, the barplot doesn't let us see where geographically these changes took place. In other words we are not able to tell if there was an increase in population in the North West of Cincinnati or the East of Cincinnati. Keeping this in mind we could have plotted a choropleth indicating the % change in each tract therefore getting more information about the change in geographical distribution of race.

In the bar plot to the left (Graph 3), we notice that almost all races in Cincinnati had a positive percentage in race except Hispanics. The population of Hispanics in Cincinnati went down almost 10% between 2000 and 2010. There was a less than 5% increase in population of african americans and almost a 140% increase in population of whites.

In the other bar plot (Graph 4), we notice that majority of the population of races in Covington had a positive %change. We notice the smallest increase in number of Hispanics between 2000 and 2010 followed by almost a 30% increase in african american population. We see the highest increase of about a 160% in population of whites.

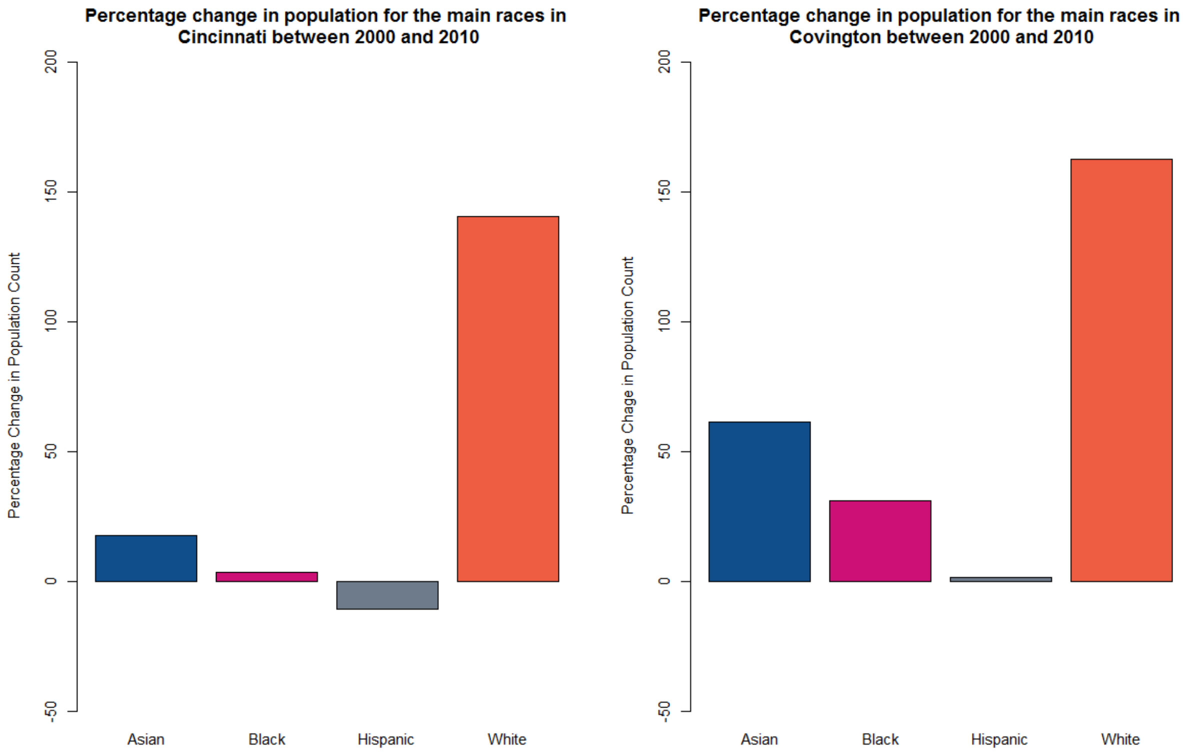



Figure 3,4 Bar plot for percentage change in population of the four major ethnic groups in Cincinnati and Covington between 2000 and 2010.

Bar plots were previously produced for the actual population counts in Cincinnati and Covington for both year 2000 and year 2010. However, due to the difference in population size in the two cities, the plots could not be produced on the same scale. In order to better compare the changes in population between the two cities, percentage change is used instead. For the old version of the graph, please refer to either website or Appendix.

**Median Income Distribution by Race in Cincinnati and Covington**

Both Graph 5 and 6 look at the relationship between income distribution and percentage of ethnic population group including Caucasian, African American, Asian and Hispanic in Cincinnati and Covington. Variables on the plot are in log form. Natural log of variables used to generate plot rather than the original variables. Because when plotting the original variables, big variations in incomes for the same ethnic population percentage are observed, and these variations make it hard to interpret trends and relationship between income and percentage of a certain ethnic group. By taking the log of both variables, we stretch the scale on both y- and x-axis to make the relationship or trend easier to see.

In Graph 5, the first plot on the top left shows an increasing trend in log median income as the log percentage of Caucasian population increases. At a higher percentage of Caucasian population that generates incomes higher than the logged mean city income, increases. The second plot on the top right shows that there is a decreasing trend between log median incomes and log percentage of African American population. As the percentage of African American population increases, more percentage of population has a log median income below the logged mean city income level. The third plot on the bottom left shows a slight increasing trend between log median incomes and the log percentage of Asian population. Most of the Asians' incomes are around the log mean income level but some of Asians' incomes are relatively lower than the mean income level. However, as the percentage of Asian population increases, the levels of income seem to be increasing. The fourth plot on the bottom right shows a decreasing trend between log median income and log percentage of Hispanic Population. As the percentage of Hispanic population increases, more people have income level below the log mean city income.

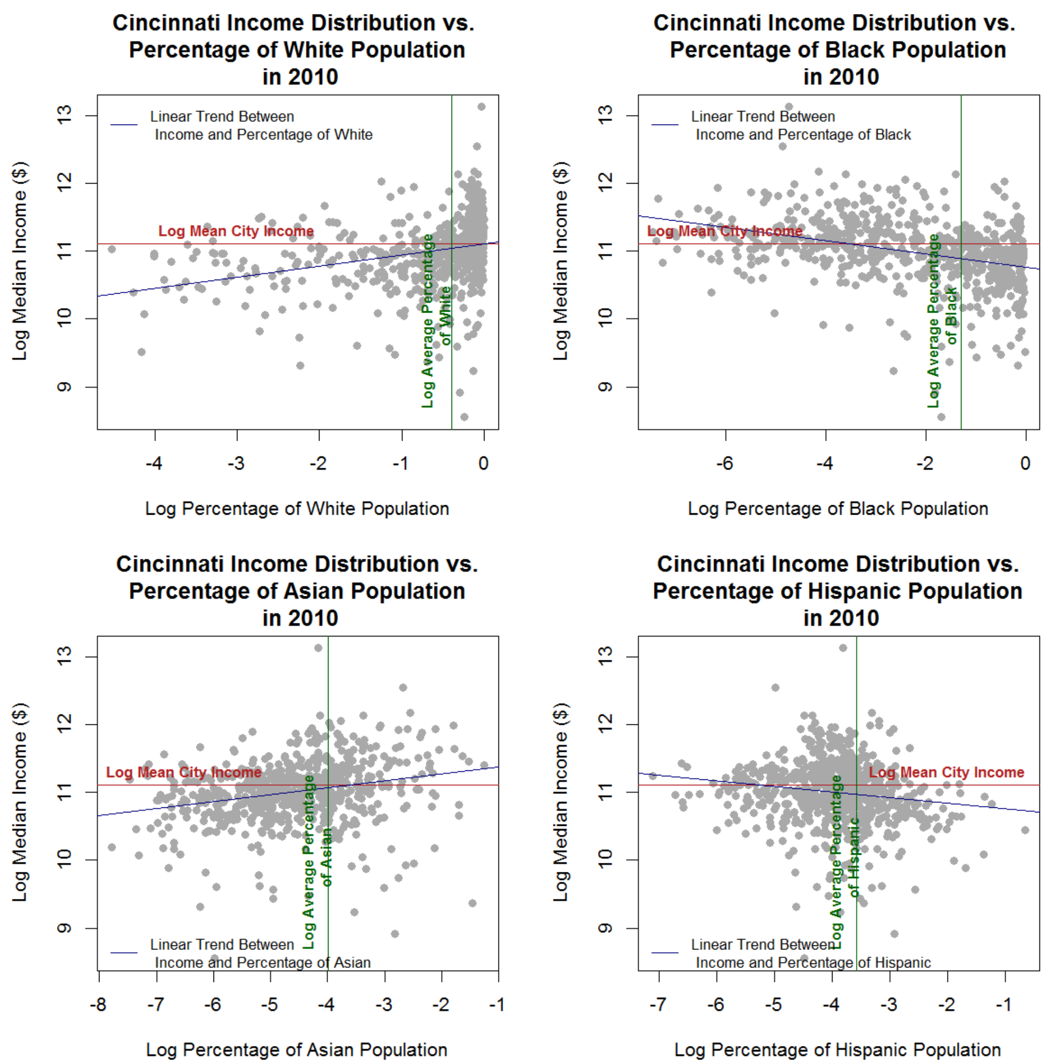



Figure 5 Scatterplot of income vs. percentage of each of the four major ethnic groups in Cincinnati. Reference lines are added to give information about the log-mean income and percentage, as well as a linear regression line showing the trend.

In Graph 6, the first plot on the top left shows an increasing trend in log median income as the log percentage of Caucasian population increases. At a higher percentage of Caucasian, more population has a higher income above the mean city income level. The second plot on the top right shows that there is a decreasing trend between log median incomes and log percentage of African American population. As the percentage of African American population increases, more population of Black has a log median income below the log mean city income level. The third plot on the bottom left does not show an apparent trend between log median incomes and the log percentage of Asian population. Most of the Asians' incomes are around the mean income level but some of Asians' incomes are relatively lower than the mean income level. The fourth plot on the bottom right shows a decreasing trend between log median income and log percentage of Hispanic population. As the percentage of Hispanic population increases, more people have income level below the log mean city income.

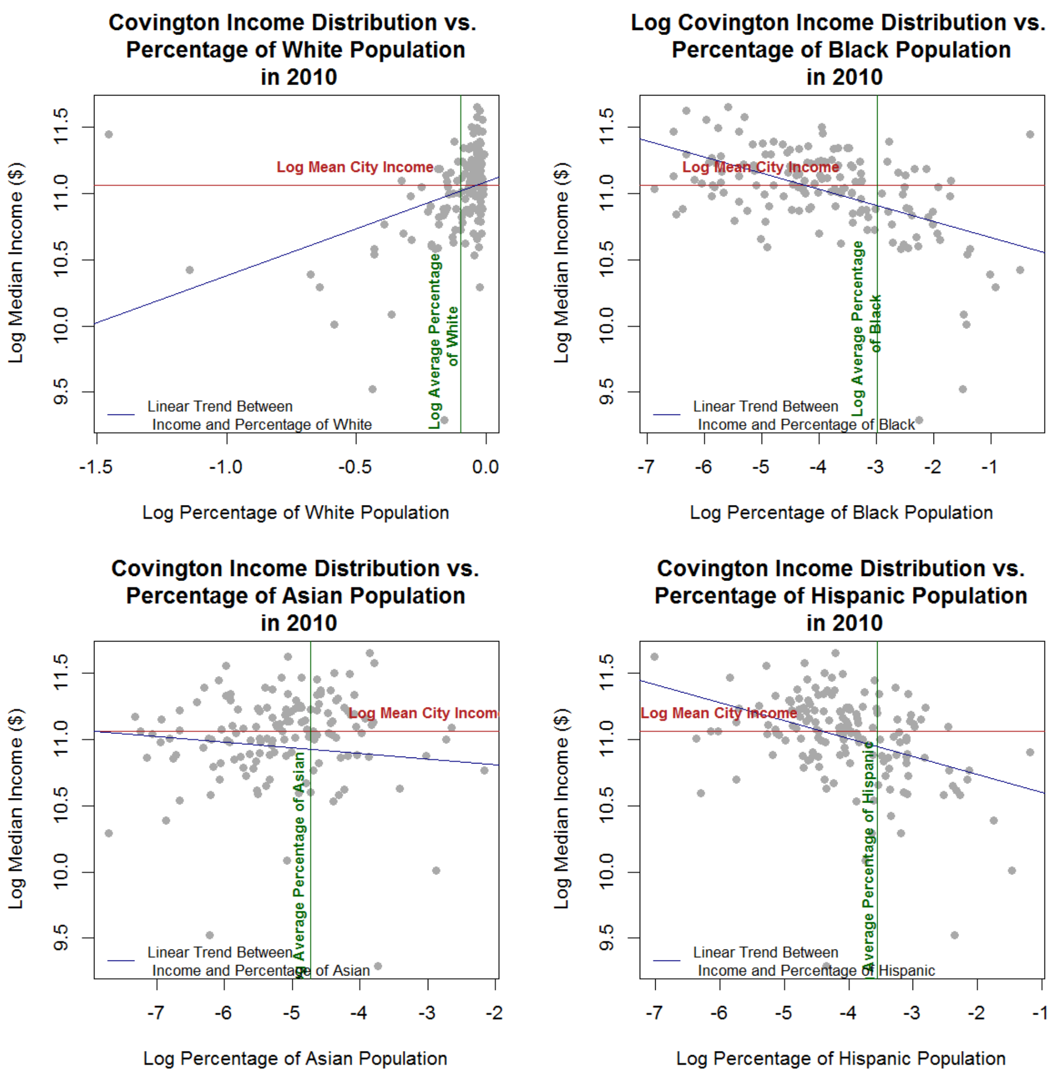



Figure 6 Scatterplot of income vs. percentage of each of the four major ethnic groups in Cincinnati. Reference lines are added to give information about the log-mean income and percentage, as well as a linear regression line showing the trend.

It should be noted that for the reference lines, we used natural log of the mean values, NOT mean of the natural log values (as an example, we used log(mean(income)) rather than mean(log(income))). This is the case because what's plotted on the scatterplots shown in Graphs 6 and 7 are the natural logged values of income and ethnic population percentage (in other words, plotting was done by taking log of each of the variables, then used as predictor and response), and the reference lines should be done in the same manner.

In Graph 7, we combine Cincinnati and Covington together with color in median income level and we see that the log median income level is higher in Covington than that in Cincinnati. We color the median income in three colors, representing the minimum, median and maximum of the median income level in both cities. Covington in general has a median income level higher than 63,980 dollars while Cincinnati has a median income level around 10,830 dollars. Since we know from the previous race distribution plots above that Covington has a larger percentage of white population than Cincinnati does, we would expect the median income in Covington is higher than that in Cincinnati due to differences in racial distribution; thus the coloring in Covington is darker in general.

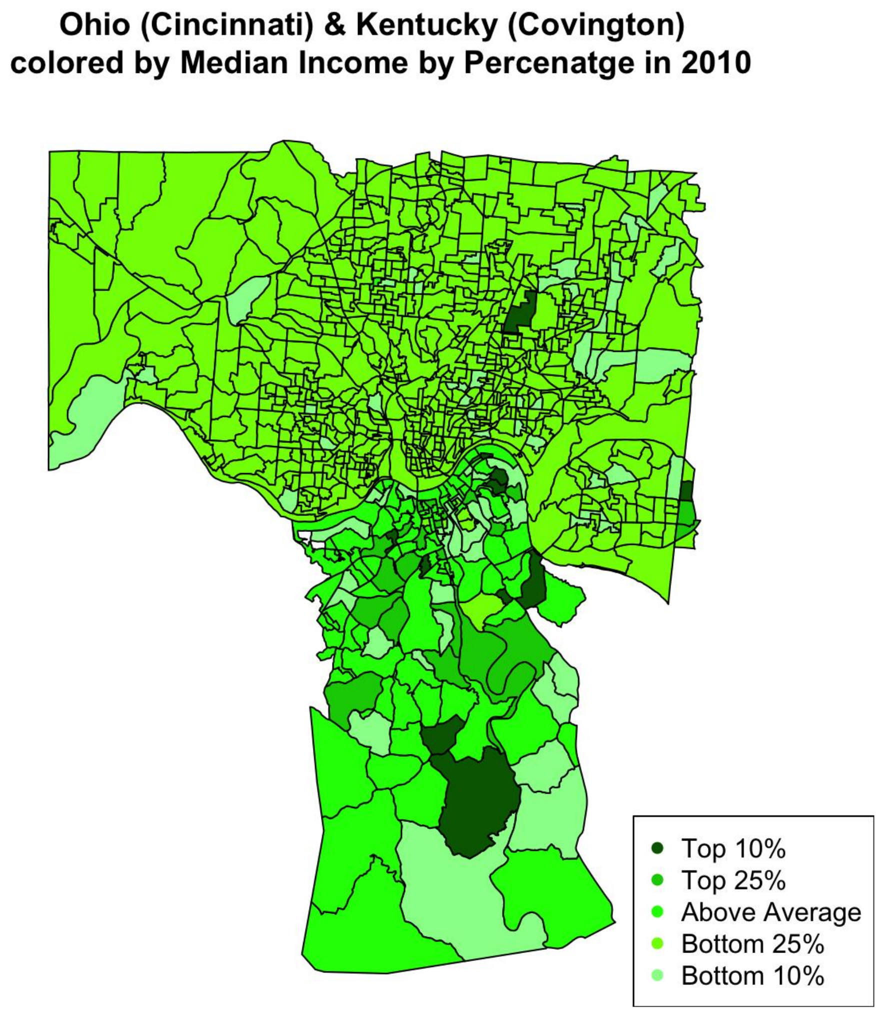



Figure 7 Choropleth of Cincinnati and Covington by income percentile (full income range combines both the tract-wide median incomes of Cincinnati and in Covington

An old version of the choropleth is included in the appendix.

**Age and Population Distribution by Gender in Cincinnati and Covington**

To analyze the age distribution, we use population pyramid using the pyramid() function from package epicalc in R. For each tract in Cincinnati and Covington, the median male and the female ages are treated as if they are the age of an individual and plotted it into a population pyramid. By visualizing, how the two pyramids are different from each other, we can get information about the population-age structure by gender for the two cities. On the x-axis is the number of tracts and on the y-axis is the group of median age; each of the age group spans 5 years. On the left-hand side is the female median age distribution and on the right-hand side is the male median age distribution. The height of each bar shows the number of tract belonging to Cincinnati or Covington that has a median age in each age group.

The population pyramid of Cincinnati shows that the largest median female age group is between 40 to 45 years old, as there are about 150 tracts in the group. The largest median male age group is between 35 to 40 years old, as there are also about 150 tracts in the group. The distributions of median age of both female and male in Cincinnati are like bell-shaped. There is less younger and elder generation in Cincinnati for both female and male. The range of median age in Cincinnati is from 5 years old to 80 years old. Compared to that of Cincinnati, the largest median age group is between 30 to 35 years old for both female and male in Covington. In addition, the range of median age spans from 5 to only 60 years old, meaning that the population in Covington in general is younger than population in Cincinnati.

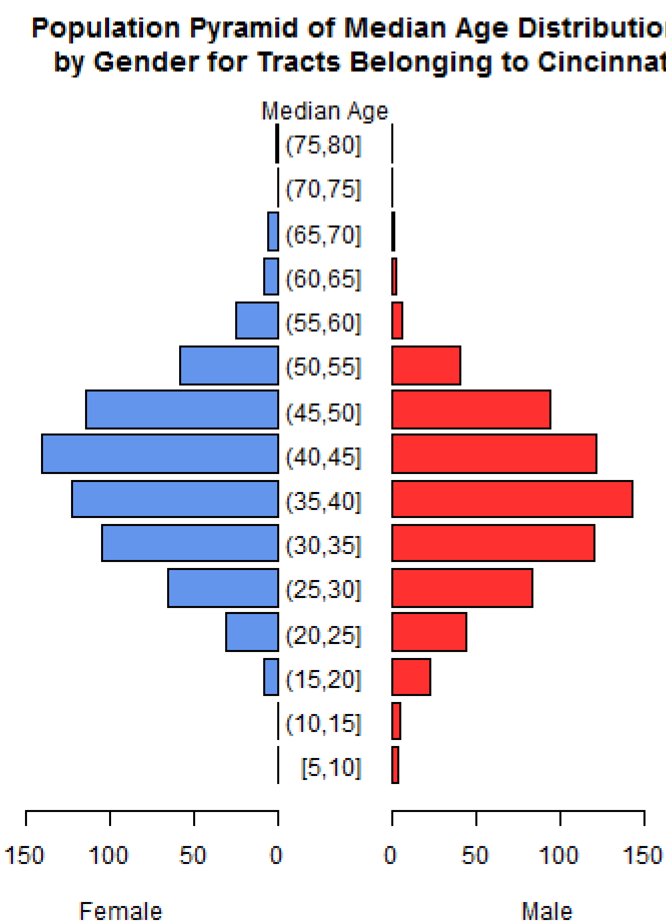



Figure 8 Population pyramid for Cincinnati.

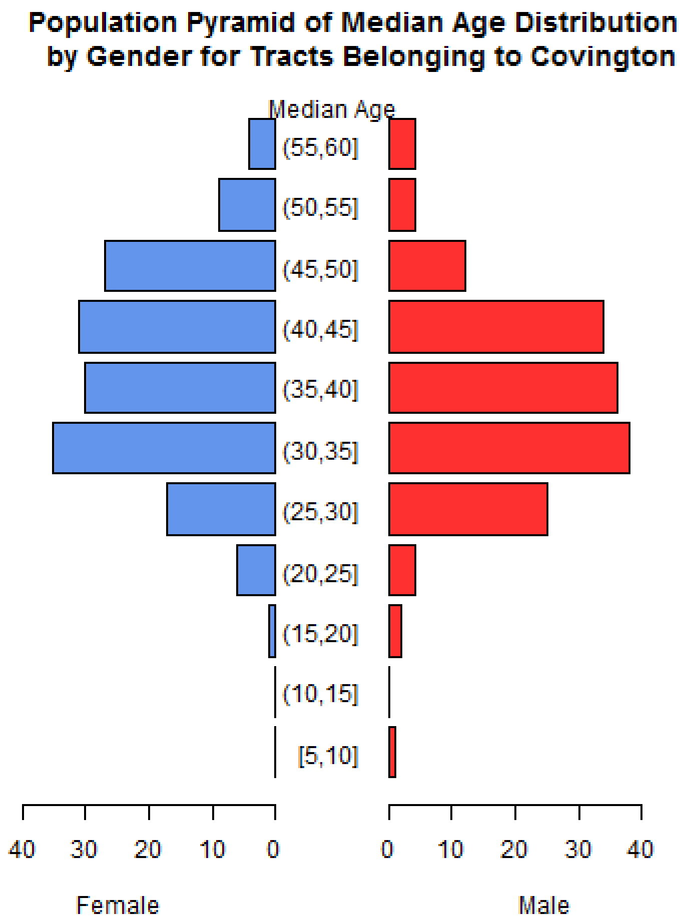



Figure 8 Population pyramid for Cincinnati.

**IMPROVEMENT**

· **Possible alternative graph options**

As alternative for the population pyramid, any of boxplot, violin plot, or bean plot can be used to show the distribution of age in Cincinnati and Covington. Population pyramid is chosen for the project because we can compare the age distribution by gender in a more direct manner (since the plots are right next to each other and share the same axis, rather than manually combining to plots together). Among the three alternatives, bean plot is of our preference, as it not only gives information about the shape of the distribution, but also explicitly marks out the median value and the intensity at each data value through the use of horizontal lines.

A combined bean plot for comparing the age distribution by gender and by city is included in the Appendix section.

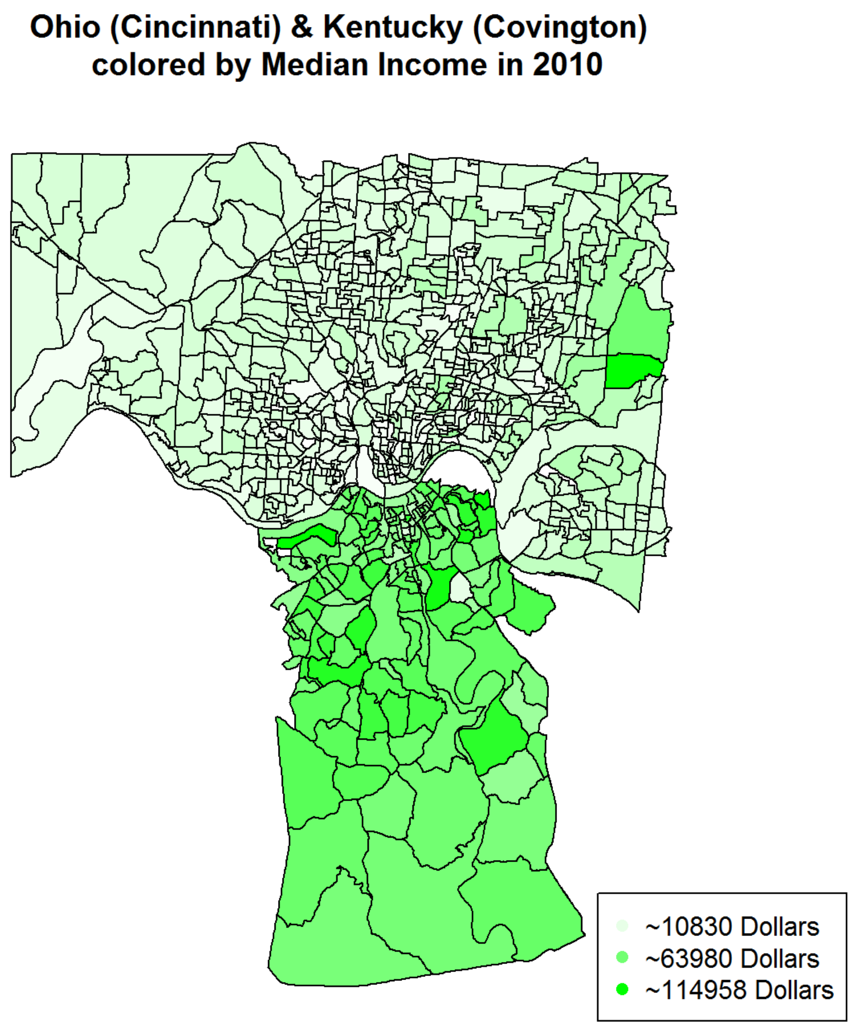
· **Possible things to explore/potential improvements if we had more time to work on the project**

It is shown in the previous sections that there are relationships between race distribution and income. One step further on this topic could be a density plot/boxplot that shows how income are distributed for

different races in two cities, and we would expect to see the incomes of African American and Hispanics to have lower values overall, comparing to income of Caucasians.
It is also of our interest to explore the distribution of income by gender, since a salary gap is known to exist between male and female workers. An example done with LOWESS density plot is given in Appendix section, plots the change in income with respect to age by gender and includes lines for both cities.

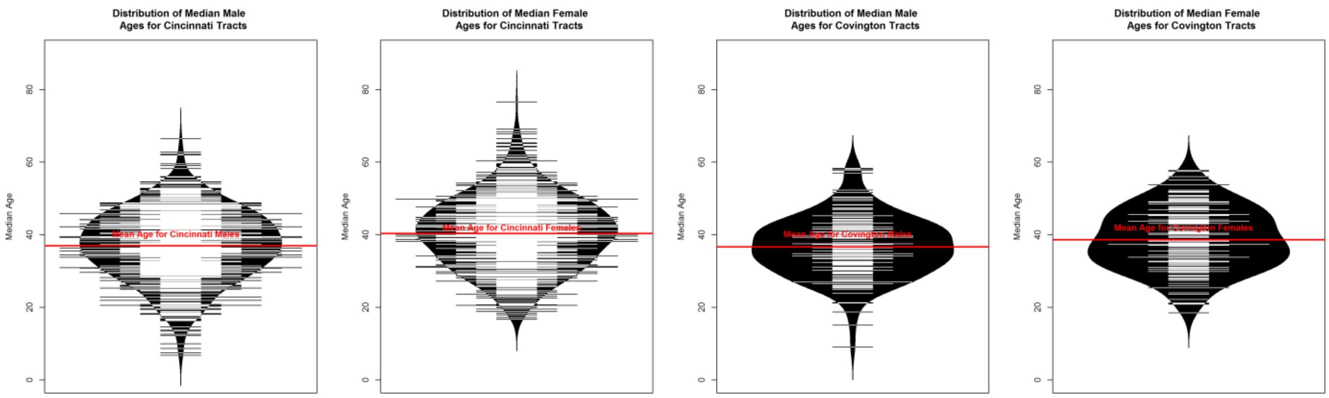
**APPENDIX**

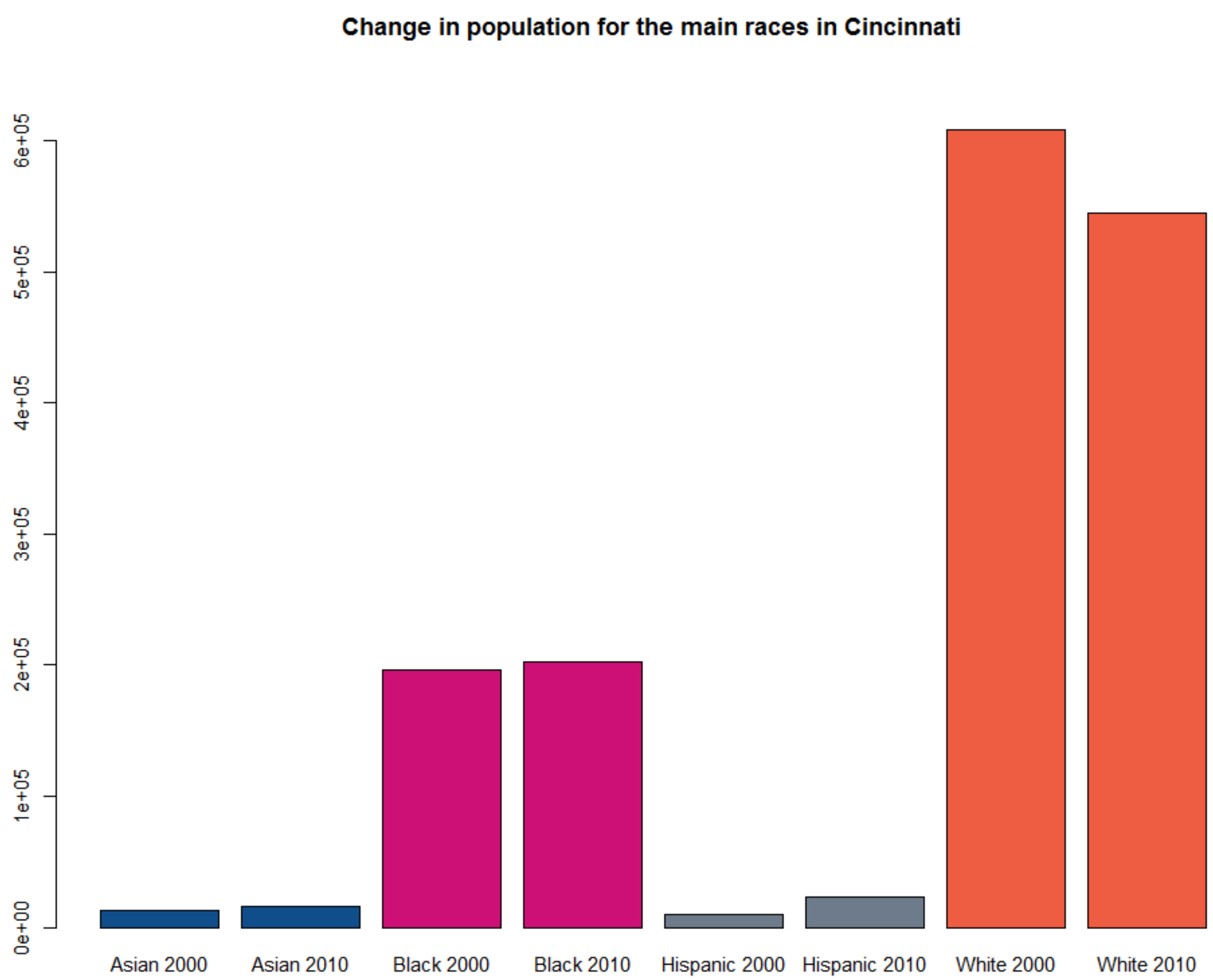
Plots:

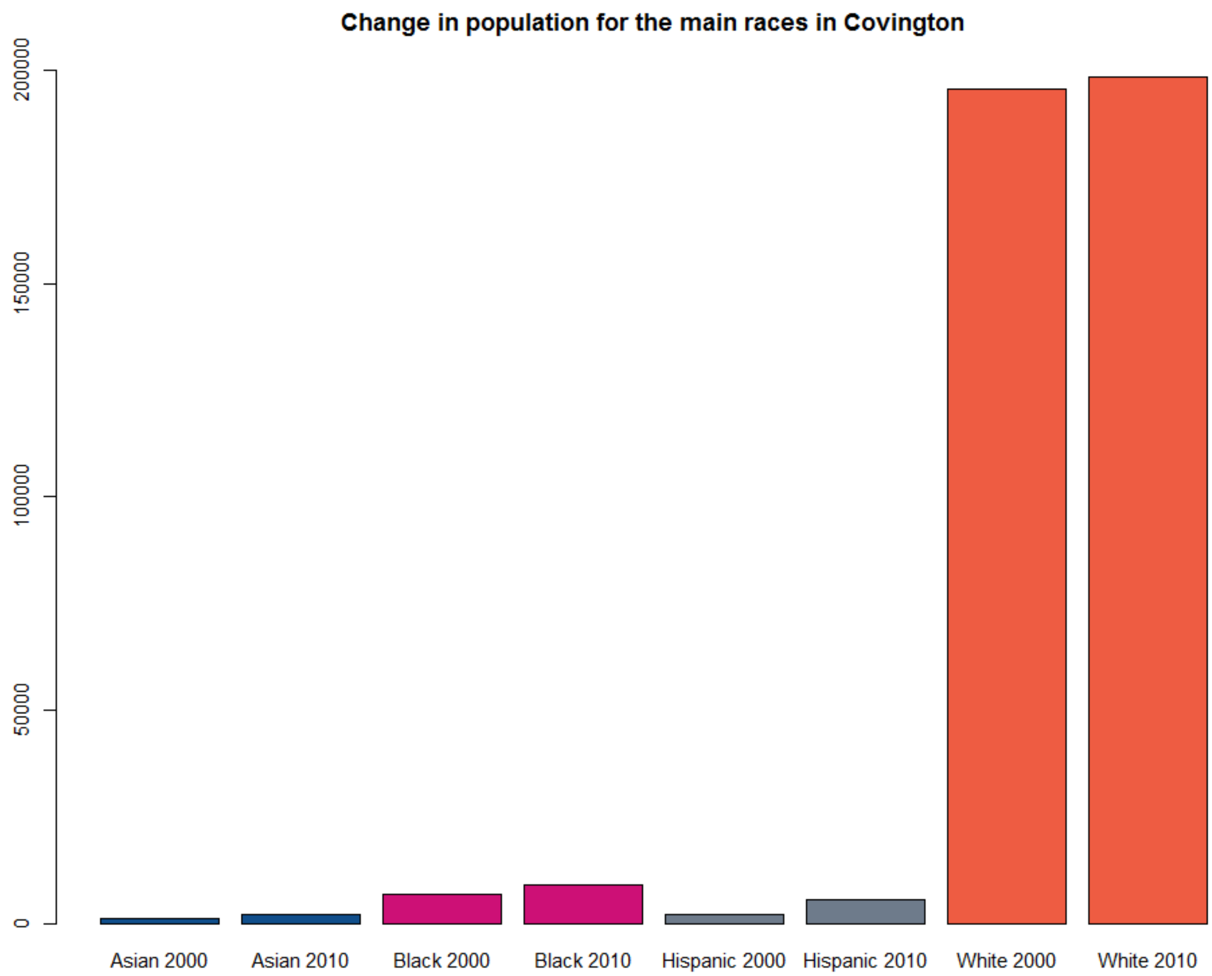



Old version of income choropleth

Combined bean plots for age distribution

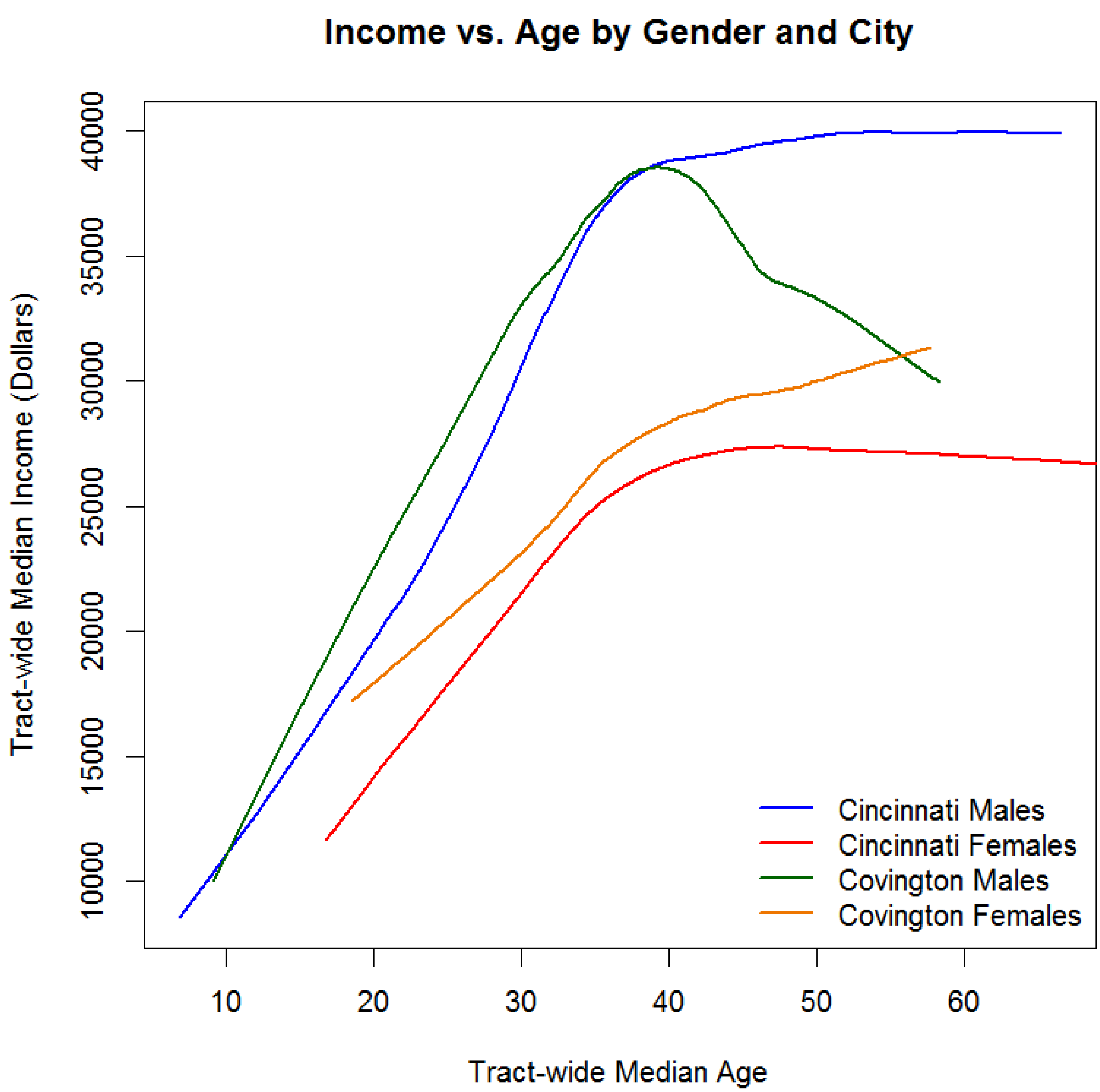



Change in population for the main races in Cincinnati
0e+00
1e+05
2e+05
3e+05
4e+05
5e+05
6e+05
Asian 2000
Asian 2010
Black 2000
Black 2010
Hispanic 2000
Hispanic 2010
White 2000
White 2010




Old versions of population bar plots

LOWESS smoothing for income by gender and age



Code used for this project:

```
#install.packages("maptools")
library(maptools)
#install.packages("raster")
##### Loading necessary libraries and packages #####
library(raster)
#install.packages("UScensus2010")
library(UScensus2010)
library(MASS)
#install.packages("Hmisc")
library(wvioplot)
#install.packages("beanplot")
library(beanplot)
#install.blkgrp("osx")  #substitute with "windows" or "osx"
library(ggplot2)
#install.packages("ggmap")
library(ggmap)
#install.packages("maps")
library(maps)

##### Data manipulation #####
#Downloads US county level data on US Census.
library("UScensus2010blkgrp")
library("UScensus2000blkgrp")
data(ohio.blkgrp)
data(ohio.blkgrp10)

##Combining income and age data from ACS with block-group data from Census libraries##
oh.acs<-read.csv("Downloads/acs-geographies/g20115oh.csv")
ac <- function(x) as.numeric(as.character(x))
name.usc <- paste(ac(ohio.blkgrp10$tract), ac(ohio.blkgrp10$blkgrp), sep="")
name.acs <- paste(ac(oh.acs[,14]), ac(oh.acs[,15]), sep="")

trials <- oh.acs[match(name.usc, name.acs), 5]

#median income
state<-"oh"
income.file <- read.csv(paste("Downloads/acs-files-quick/e20115",state,"0064000.txt",sep=""), header=FALSE)

rows<-match(trials,income.file[,6])
ohio.blkgrp10$income.male <- ac(income.file[rows, 51])
ohio.blkgrp10$income.female <- ac(income.file[rows, 52])

age.file <- read.csv(paste("Downloads/acs-files-quick/e20115",state,"0003000.txt",sep=""), header=FALSE)
```

```
rows <- match(trials, age.file[,6])
ohio.blkgrp10$age.male <- ac(age.file[rows, 101])
ohio.blkgrp10$age.female <- ac(age.file[rows, 102])
par(mfrow=c(2,2))
hist(ohio.blkgrp10$income.male)
hist(ohio.blkgrp10$income.female)
hist(ohio.blkgrp10$age.male)
hist(ohio.blkgrp10$age.female)

#Repeat for Kentucky
data(kentucky.blkgrp)
data(kentucky.blkgrp10)
ky.acs<-read.csv("Downloads/acs-geographies/g20115ky.csv")
name.usc <- paste(ac(kentucky.blkgrp10$tract), ac(kentucky.blkgrp10$blkgrp), sep="")
name.acs <- paste(ac(ky.acs[,14]), ac(ky.acs[,15]), sep="")

trials <- ky.acs[match(name.usc, name.acs), 5]

#median income
state<-"ky"
income.file <- read.csv(paste("Downloads/acs-files-quick/e20115",state,"0064000.txt",sep=""), header=FALSE)

rows<-match(trials,income.file[,6])
kentucky.blkgrp10$income.male <- ac(income.file[rows, 51])
kentucky.blkgrp10$income.female <- ac(income.file[rows, 52])

age.file <- read.csv(paste("Downloads/acs-files-quick/e20115",state,"0003000.txt",sep=""), header=FALSE)
rows <- match(trials, age.file[,6])
kentucky.blkgrp10$age.male <- ac(age.file[rows, 101])
kentucky.blkgrp10$age.female <- ac(age.file[rows, 102])

par(mfrow=c(2,2))
hist(kentucky.blkgrp10$income.male)
hist(kentucky.blkgrp10$income.female)
hist(kentucky.blkgrp10$age.male)
hist(kentucky.blkgrp10$age.female)


#Now we have two spactial polygon objects for Ohio and Kentucky, respectively.
#Before moving on to analyzing the population distribution, race distribution,
#and their relationship with income and age, we need to extract the tracts
#correspond to our cities of interest: 1) Cincinnati (Ohio) and 2) Covington (Kentucky)

#Get cincinnati tracts
```

```
cincinnati<-vector()
#Using a for loop that goes through each data row in the Ohio spacial polygon,
#we pull out all row number for tracts whose geographical coordinates are within the range
#corresponding to Cincinnati
for(i in 1:nrow(ohio.blkgrp)){
  curr.cor<-coordinates(ohio.blkgrp[i,])
  if(curr.cor[1]<(-84.3) & curr.cor[2]<(39.3)){
    cincinnati<-append(cincinnati,i)
  }
}
#Subsetting into a Cincinnati spatial polygon using the row number obtained above
cincinnati.tracts<-ohio.blkgrp[cincinnati,]

#Plot and check
png("cincinnati_tracts_check.png")
plot(cincinnati.tracts)
dev.off()

#Repeat for the 2010 polygon
cincinnati10<-vector()
for(i in 1:nrow(ohio.blkgrp10)){
  curr.cor<-coordinates(ohio.blkgrp10[i,])
  if(curr.cor[1]<(-84.3) & curr.cor[2]<(39.3)){
    cincinnati10<-append(cincinnati10,i)
  }
}
cincinnati.tracts10<-ohio.blkgrp10[cincinnati10,]
png("Cincinnati_tracts10_check.png")
plot(cincinnati.tracts10)
dev.off()

#Get Covington tracts
covington<-vector()
#Using a for loop that goes through each data row in the Kentucky spacial polygon,
#we pull out all row number for tracts whose geographical coordinates are within the range
#corresponding to Covington
for(i in 1:nrow(kentucky.blkgrp)){
  curr.cor<-coordinates(kentucky.blkgrp[i,])
  if(curr.cor[1]>(-84.6) & curr.cor[1]<(-84.4) & curr.cor[2]>(38.80)){
    covington<-append(covington,i)
  }
}

#Subsetting
```

```
covington.tracts<-kentucky.blkgrp[covington,]
#Plot and check
png("Covington_tracts_check.png")
plot(covington.tracts)
dev.off()

#Repeat for the Census 2010 Polygon
covington10<-vector()
for(i in 1:nrow(kentucky.blkgrp10)){
  curr.cor<-coordinates(kentucky.blkgrp10[i,])
  if(curr.cor[1]>(-84.6) & curr.cor[1]<(-84.4) & curr.cor[2]>(38.80)){
    covington10<-append(covington10,i)
  }
}
covington.tracts10<-kentucky.blkgrp10[covington10,]
png("Covington_tracts10_check.png")
plot(covington.tracts10)
dev.off()

#Data frame combine and extration are done. Start to analyze the data.
#####################################

##### Population distribution by race #####
#Barplots: How are population(distribution) of the four major ethnic groups differ
#for Cincinnati and Covington

#Appeared in presentation- plotting barplots of acutal population counts for major
#ethnic groups and horizontally combine plot of 2000 and 2010.
png("Barplot Cin Population by race.png",width=1000,height=800,res=90)
barplot(c(sum(cincinnati.tracts$asian),sum(cincinnati.tracts10$P0080006),sum(cincinnati.tracts$black),
          sum(cincinnati.tracts10$P0080004),sum(cincinnati.tracts$hispanic),sum(cincinnati.tracts10$P0070009),
        sum(cincinnati.tracts$white),sum(cincinnati.tracts10$P0080003)),
        names.arg=c("Asian 2000","Asian 2010","Black 2000","Black 2010","Hispanic 2000","Hispanic 
2010","White 2000","White 2010"),
        main="Change in population for the main races in Cincinnati",

col=c("dodgerblue4","dodgerblue4","deeppink3","deeppink3","lightsteelblue4","lightsteelblue4","tomato2","toma
to2"),
        ylim=c(0,6.5e+05))
dev.off()

png("Barplot Cov Population by race.png",width=1000,height=800,res=90)
barplot(c(sum(covington.tracts$asian),sum(covington.tracts10$P0080006),sum(covington.tracts$black),
          sum(covington.tracts10$P0080004),sum(covington.tracts$hispanic),sum(covington.tracts10$P0070009),
```

```
     sum(covington.tracts$white),sum(covington.tracts10$P0080003)),
    names.arg=c("Asian 2000","Asian 2010","Black 2000","Black 2010","Hispanic 2000","Hispanic
2010","White 2000",
          "White 2010"),main="Change in population for the main races in Covington",

col=c("dodgerblue4","dodgerblue4","deeppink3","deeppink3","lightsteelblue4","lightsteelblue4","tomato2","toma
to2")
    ,ylim=c(0,200000))
dev.off()


#Improved version done after presentation- instead of plotting the actual population,
#we plot the changes, so that the values for Cincinnati and Covington are on the same scale

changeAsianCin<-(sum(cincinnati.tracts10$P0080006)-
sum(cincinnati.tracts$asian))/sum(cincinnati.tracts$asian)*100
changeBlackCin<-(sum(cincinnati.tracts10$P0080004)-
sum(cincinnati.tracts$black))/sum(cincinnati.tracts$black)*100
changeWhiteCin<-(sum(cincinnati.tracts10$P0080003)-
sum(cincinnati.tracts$white))/sum(cincinnati.tracts$white)*100
changeHispanicCin<-(sum(cincinnati.tracts10$P0070009)-
sum(cincinnati.tracts$hispanic))/sum(cincinnati.tracts$hispanic)*100
changeAsianCov<-(sum(covington.tracts10$P0080006)-
sum(covington.tracts$asian))/sum(covington.tracts$asian)*100
changeBlackCov<-(sum(covington.tracts10$P0080004)-
sum(covington.tracts$black))/sum(covington.tracts$black)*100
changeWhiteCov<-(sum(covington.tracts10$P0080003)-
sum(covington.tracts$white))/sum(covington.tracts$white)*100
changeHispanicCov<-(sum(covington.tracts10$P0070009)-
sum(covington.tracts$hispanic))/sum(covington.tracts$hispanic)*100

png("Barplot percentage change in population by race.png",width=1200,height=800,res=90)
par(mfrow=c(1,2))
barplot(c(changeAsianCin,changeBlackCin,changeWhiteCin,changeHispanicCin),names.arg=c("Asian","Black","
Hispanic","White"),
    main="Percentage change in population for the main races in\n Cincinnati between 2000 and 2010",
    col=c("dodgerblue4","deeppink3","lightsteelblue4","tomato2"),
    ylab="Percentage Change in Population Count",ylim=c(-50,200))
barplot(c(changeAsianCov,changeBlackCov,changeWhiteCov,changeHispanicCov),names.arg=c("Asian","Black"
,"Hispanic","White"),
    main="Percentage change in population for the main races in\n Covington between 2000 and 2010",
    col=c("dodgerblue4","deeppink3","lightsteelblue4","tomato2"),
    ylab="Percentage Chage in Population Count",ylim=c(-50,200))
dev.off()
```

```
#choropleth: Visualize the change in race distribution from 2000 to 2010, Cincinnati and Covington together
#plot(ohio.blkgrp10)
png("Choropleth Cin & Cov White Percent.png",width=2000,height=2000,res=200)
par(mfrow=c(1,2))
summary(c(percentage.whites.cin,percentage.whites.cov))
percentage.whites.cin<-c((cincinnati.tracts$white/cincinnati.tracts$pop2000),
                        (cincinnati.tracts10$P0080003/cincinnati.tracts10$P0010001))
cols.white.cin<-ifelse(percentage.whites.cin>0.9,"darkseagreen4",
                      ifelse(percentage.whites.cin<0.5,"darkseagreen1","darkseagreen3"))
percentage.whites.cov<-c((covington.tracts$white/covington.tracts$pop2000),
                        (covington.tracts10$P0080003/covington.tracts10$P0010001))
cols.white.cov<-
ifelse(percentage.whites.cov>0.9,"darkseagreen4",ifelse(percentage.whites.cov<0.5,"darkseagreen1","darkseagree
n3"))

cols.white.cin2000<-cols.white.cin[1:nrow(cincinnati.tracts)]
cols.white.cov2000<-cols.white.cov[1:nrow(covington.tracts)]

border2000<-c(rep("brown4",nrow(cincinnati.tracts)),rep("dodgerblue4",nrow(covington.tracts)))
#Choropleth of Cincinnati and Covington by percentage of white population: 2000 vs.2010
plot(rbind(cincinnati.tracts,covington.tracts),col=c(cols.white.cin2000,cols.white.cov2000),border=border2000)
title("Ohio (Cincinnati) & Kentucky (Covington) \n Colored by Percentage of Whites in 2000")
legend("bottomleft",c("More than 90%","More than 50%","Less than 50%"),
       fill=c("darkseagreen4","darkseagreen3","darkseagreen1"))

cols.white.cin2010<-cols.white.cin[(nrow(cincinnati.tracts)+1):length(cols.white.cin)]
cols.white.cov2010<-cols.white.cov[(nrow(covington.tracts)+1):length(cols.white.cov)]

border2010<-c(rep("brown4",nrow(cincinnati.tracts10)),rep("dodgerblue4",nrow(covington.tracts10)))

plot(rbind(cincinnati.tracts10,covington.tracts10),col=c(cols.white.cin2010,cols.white.cov2010),border=border20
10)
title("Ohio (Cincinnati) & Kentucky (Covington) \n Colored by Percentage of Whites in 2010")
legend("bottomleft",c("More than 90%","More than 50%","Less than 50%"),
       fill=c("darkseagreen4","darkseagreen3","darkseagreen1"))

dev.off()

#Repeat the Choropleth, now for African American
png("Choropleth Cin & Cov Black Percent.png",width=2000,height=2000,res=200)
par(mfrow=c(1,2))
percentage.blacks.cin<-
c((cincinnati.tracts$black/cincinnati.tracts$pop2000),(cincinnati.tracts10$P0080004/cincinnati.tracts10$P0010001
))
```

```
cols.black.cin<-
ifelse(percentage.blacks.cin>0.35,"darkseagreen4",ifelse(percentage.blacks.cin<0.05,"darkseagreen1","darkseagre
en3"))
percentage.blacks.cov<-
c((covington.tracts$black/covington.tracts$pop2000),(covington.tracts10$P0080004/covington.tracts10$P001000
1))
cols.black.cov<-
ifelse(percentage.blacks.cov>0.35,"darkseagreen4",ifelse(percentage.blacks.cov<0.05,"darkseagreen1","darkseagr
een3"))
summary(c(percentage.blacks.cin,percentage.blacks.cov))

cols.black.cin2000<-cols.black.cin[1:nrow(cincinnati.tracts)]
cols.black.cov2000<-cols.black.cov[1:nrow(covington.tracts)]

border2000<-c(rep("brown4",nrow(cincinnati.tracts)),rep("dodgerblue4",nrow(covington.tracts)))

plot(rbind(cincinnati.tracts,covington.tracts),col=c(cols.black.cin2000,cols.black.cov2000),border=border2000)
title("Ohio (Cincinnati) & Kentucky (Covington) \n colored by Percentage of Blacks in 2000")
legend("bottomleft",c("Less than 5%","Less than 35%","More than
35%"),fill=c("darkseagreen1","darkseagreen3","darkseagreen4"))

cols.black.cin2010<-cols.black.cin[(nrow(cincinnati.tracts)+1):length(cols.black.cin)]
cols.black.cov2010<-cols.black.cov[(nrow(covington.tracts)+1):length(cols.black.cov)]

border2010<-c(rep("brown4",nrow(cincinnati.tracts10)),rep("dodgerblue4",nrow(covington.tracts10)))

plot(rbind(cincinnati.tracts10,covington.tracts10),col=c(cols.black.cin2010,cols.black.cov2010),border=border20
10)
title("Ohio (Cincinnati) & Kentucky (Covington) \n colored by Percentage of Blacks 2010")
legend("bottomleft",c("Less than 5%","Less than 35%","More than
35%"),fill=c("darkseagreen1","darkseagreen3","darkseagreen4"))

dev.off()

######for final proj you only need code up till here, go on to next section#####
#Default method:
choropleth(ohio.blkgrp10)
ohio.blkgrp10$P0030001
choropleth(ohio.blkgrp10$P0030001)

head(ohio.blkgrp10)

#
install.packages("UScensus2000blkgrp")
```

```
data(ohio.blkgrp)
dim(ohio.blkgrp)
dim(ohio.blkgrp10)
install.packages("UScensus2000tract")
library("UScensus2000tract")
data(ohio.tract)
dim(ohio.tract)

rownames(ohio.tract)
head(ohio.blkgrp)

summary(ohio.blkgrp$white)

choropleth(ohio.blkgrp$city)
choropleth(ohio.blkgrp10)

color.grad <- function(input){
  sapply(input, function(kk){
    if(is.na(kk)){return(1)}
    else{rgb(1-kk/max(input,na.rm=TRUE), 1-kk/max(input,na.rm=TRUE), 1)}
  }
)}
par(mfrow=c(1,2))
cols.white<-color.grad(ohio.blkgrp$white)
plot(ohio.blkgrp,col=cols.white,main="Ohio colored by Number of whites",xlim=c(-84.8,-
83.8),ylim=c(38.8,39.7))
title("Ohio (Cincinnati) colored by Number of Whites 2000")
cols.white10<-color.grad(ohio.blkgrp10$P0080003)
plot(ohio.blkgrp10,col=cols.white10,main="Ohio colored by Number of whites",xlim=c(-84.8,-
83.8),ylim=c(38.8,39.7))
title("Ohio (Cincinnati) colored by Number of Whites 2010")
par(mfrow=c(1,2))
cols.black<-color.grad(ohio.blkgrp$black)
plot(ohio.blkgrp,col=cols.black,main="Ohio colored by Number of blacks",xlim=c(-84.8,-83.8),ylim=c(38.8,39.7))
title("Ohio (Cincinnati) colored by Number of Blacks 2000")
cols.black10<-color.grad(ohio.blkgrp10$P0080004)
plot(ohio.blkgrp10,col=cols.black10,main="Ohio colored by Number of blacks",xlim=c(-84.8,-
83.8),ylim=c(38.8,39.7))
title("Ohio (Cincinnati) colored by Number of Blacks 2010")
par(mfrow=c(1,2))
cols.asian<-color.grad(ohio.blkgrp$asian)
plot(ohio.blkgrp,col=cols.asian,main="Ohio colored by Number of Asians",xlim=c(-84.8,-83.8),ylim=c(38.8,39.7))
title("Ohio (Cincinnati) colored by Number of Asians 2000")
cols.asian10<-color.grad(ohio.blkgrp10$P0080006)
```

```
plot(ohio.blkgrp10,col=cols.asian10,main="Ohio colored by Number of Asians",xlim=c(-84.8,-83.8),ylim=c(38.8,39.7))
title("Ohio (Cincinnati) colored by Number of Asians 2010")

color.grad <- function(input){
  sapply(input, function(kk){
    if(is.na(kk)){return(1)}
    else{rgb(1-kk/max(input,na.rm=TRUE), 1-kk/max(input,na.rm=TRUE), 1)}
  }
  )}

#Population shares for African-American
blackohio2010<-ohio.blkgrp10$P0080004
blackohio2000<-
cols.black10<-color.grad(ohio.blkgrp10$P0080004)
plot(ohio.blkgrp10,col=cols.black10 ,xlim=c(-84.8,-83.8),ylim=c(38.8,39.7))
title("Population of Afriac-American in Cincinnati in 2010")


cols.his10<-color.grad(ohio.blkgrp10$P0080008)
plot(ohio.blkgrp10,col=cols.his10 ,xlim=c(-84.8,-83.8),ylim=c(38.8,39.7))
title("Population of Hispanic in Cincinnati in 2010")



##### Income and age, and their correlation with population distribution #####

#Income by race: Scatterplot of tract-wide median income by percent of ethnic population
#with reference lines added for city-wide median income and average percentage of ethnic
#group of interest

#Note that for the purpose of including a linear regression trend line, after calculating
#the percentage of each ethinic group in a certain tract, we need to replace the -Inf
#values (due to 0 input) with NA. We do it only for the linear trend part, so the
#replacements do not affect the overall look of our plots.
#Cincinnati
income.cin<-cincinnati.tracts10$income.male+cincinnati.tracts10$income.female
white.ave.cin<-mean(cincinnati.tracts10$P0080003/cincinnati.tracts10$P0010001)
black.ave.cin<-mean(cincinnati.tracts10$P0080004/cincinnati.tracts10$P0010001)
asian.ave.cin<-mean(cincinnati.tracts10$P0080006/cincinnati.tracts10$P0010001)
hispanic.ave.cin<-mean(cincinnati.tracts10$P0070009/cincinnati.tracts10$P0010001)

png("Scatterplot Log Income by Race in Cincinnati.png",width=1200,height=1200,res=150)
par(mfrow=c(2,2))
```

```
plot(log(cincinnati.tracts10$P0080003/cincinnati.tracts10$P0010001),log(income.cin),pch=16,col="darkgrey",
     main="Cincinnati Income Distribution vs. \n Percentage of White Population \n in 2010",
     xlab="Log Percentage of White Population",ylab="Log Median Income ($)")
abline(h=log(mean(income.cin,na.rm=TRUE)),col="firebrick",lwd=1.2)
abline(v=log(white.ave.cin),col="darkgreen",lwd=1.2)
log.percent.wh.cin = log(cincinnati.tracts10$P0080003/cincinnati.tracts10$P0010001)
log.percent.wh.cin[which(log.percent.wh.cin==-Inf)]=NA
abline(lm(log(income.cin)~log.percent.wh.cin),col="navy",lwd=1.2)
text(-3,11.3, "Log Mean City Income", col = "firebrick",cex=0.8,font=2)
text(-0.6,10, "Log Average Percentage \n of White",col = "darkgreen", srt=90,cex=0.8,font=2)
legend("topleft","Linear Trend Between \n Income and Percentage of White",col="navy",lwd=1.2,
       cex=0.8,bty="n")


plot(log(cincinnati.tracts10$P0080004/cincinnati.tracts10$P0010001),log(income.cin),pch=16,col="darkgrey",
     main="Cincinnati Income Distribution vs. \n Percentage of Black Population \n in 2010",
     xlab="Log Percentage of Black Population",ylab="Log Median Income ($)")
abline(h=log(mean(income.cin,na.rm=TRUE)),col="firebrick",lwd=1.2)
abline(v=log(black.ave.cin),col="darkgreen",lwd=1.2)
log.percent.bl.cin = log(cincinnati.tracts10$P0080004/cincinnati.tracts10$P0010001)
log.percent.bl.cin[which(log.percent.bl.cin==-Inf)]=NA
abline(lm(log(income.cin)~log.percent.bl.cin),col="navy",lwd=1.2)
text(-6,11.3, "Log Mean City Income", col = "firebrick",cex=0.8,font=2)
text(-1.7,10, "Log Average Percentage \n of Black",col = "darkgreen", srt=90,cex=0.8,font=2)
legend("topleft","Linear Trend Between \n Income and Percentage of Black",col="navy",lwd=1.2,
       cex=0.8,bty="n")

plot(log(cincinnati.tracts10$P0080006/cincinnati.tracts10$P0010001),log(income.cin),pch=16,col="darkgrey",
     main="Cincinnati Income Distribution vs.\n Percentage of Asian Population \n in 2010",
     xlab="Log Percentage of Asian Population",ylab="Log Median Income ($)")
abline(h=log(mean(income.cin,na.rm=TRUE)),col="firebrick",lwd=1.2)
abline(v=log(asian.ave.cin),col="darkgreen",lwd=1.2)
log.percent.as.cin = log(cincinnati.tracts10$P0080006/cincinnati.tracts10$P0010001)
log.percent.as.cin[which(log.percent.as.cin==-Inf)]=NA
abline(lm(log(income.cin)~log.percent.as.cin),col="navy",lwd=1.2)
text(-6.5,11.3, "Log Mean City Income", col = "firebrick",cex=0.8,font=2)
text(-4.2,10, "Log Average Percentage \n of Asian",col = "darkgreen", srt=90,cex=0.8,font=2)
legend("bottomleft","Linear Trend Between \n Income and Percentage of
Asian",col="navy",lwd=1.2,cex=0.8,bty="n")


plot(log(cincinnati.tracts10$P0070009/cincinnati.tracts10$P0010001),log(income.cin),pch=16,col="darkgrey",
     main="Cincinnati Income Distribution vs.\n Percentage of Hispanic Population \n in 2010",
```

```
    xlab="Log Percentage of Hispanic Population",ylab="Log Median Income ($)")
abline(h=log(mean(income.cin,na.rm=TRUE)),col="firebrick",lwd=1.2)
abline(v=log(hispanic.ave.cin),col="darkgreen",lwd=1.2)
log.percent.hi.cin = log(cincinnati.tracts10$P0070009/cincinnati.tracts10$P0010001)
log.percent.hi.cin[which(log.percent.hi.cin==-Inf)]=NA
abline(lm(log(income.cin)~log.percent.hi.cin),col="navy",lwd=1.2)
text(-2,11.3, "Log Mean City Income", col = "firebrick",cex=0.8,font=2)
text(-3.8,10, "Log Average Percentage \n of Hispanic",col = "darkgreen", srt=90,cex=0.8,font=2)
legend("bottomleft","Linear Trend Between \n Income and Percentage of Hispanic",
      col="navy",lwd=1.2,cex=0.8,bty="n")

dev.off()

#Covington
income.cov<-covington.tracts10$income.male+covington.tracts10$income.female
white.ave.cov<-mean(covington.tracts10$P0080003/covington.tracts10$P0010001)
black.ave.cov<-mean(covington.tracts10$P0080004/covington.tracts10$P0010001)
asian.ave.cov<-mean(covington.tracts10$P0080006/covington.tracts10$P0010001)
hispanic.ave.cov<-mean(covington.tracts10$P0070009/covington.tracts10$P0010001)

png("Scatter Plot Log Income by Race in Covington.png",width=1200,height=1200,res=150)
par(mfrow=c(2,2))
plot(log(covington.tracts10$P0080003/covington.tracts10$P0010001),log(income.cov),pch=16,col="darkgrey",
    main="Covington Income Distribution vs. \n Percentage of White Population \n in 2010",
    xlab="Log Percentage of White Population",ylab="Log Median Income ($)")
abline(h=log(mean(income.cov,na.rm=TRUE)),col="firebrick",lwd=1.2)
abline(v=log(white.ave.cov),col="darkgreen",lwd=1.2)
log.percent.wh.cov = log(covington.tracts10$P0080003/covington.tracts10$P0010001)
log.percent.wh.cov[which(log.percent.wh.cov==-Inf)]=NA
abline(lm(log(income.cov)~log.percent.wh.cov),col="navy",lwd=1.2)
text(-0.5,11.2, "Log Mean City Income", col = "firebrick",cex=0.8,font=2)
text(-0.17,9.9, "Log Average Percentage \n of White",col = "darkgreen", srt=90,cex=0.8,font=2)
legend("bottomleft","Linear Trend Between \n Income and Percentage of White",
      col="navy",lwd=1.2,cex=0.8,bty="n")

plot(log(covington.tracts10$P0080004/covington.tracts10$P0010001),log(income.cov),pch=16,col="darkgrey",
    main="Log Covington Income Distribution vs. \n Percentage of Black Population \n in 2010",
    xlab="Log Percentage of Black Population",ylab="Log Median Income ($)")
abline(h=log(mean(income.cov,na.rm=TRUE)),col="firebrick",lwd=1.2)
abline(v=log(black.ave.cov),col="darkgreen",lwd=1.2)
log.percent.bl.cov = log(covington.tracts10$P0080004/covington.tracts10$P0010001)
log.percent.bl.cov[which(log.percent.bl.cov==-Inf)]=NA
abline(lm(log(income.cov)~log.percent.bl.cov),col="navy",lwd=1.2)
text(-5,11.2, "Log Mean City Income", col = "firebrick",cex=0.8,font=2)
```

```
text(-3.2,10, "Log Average Percentage \n of Black",col = "darkgreen", srt=90,cex=0.8,font=2)
legend("bottomleft","Linear Trend Between \n Income and Percentage of Black",
       col="navy",lwd=1.2,cex=0.8,bty="n")

plot(log(covington.tracts10$P0080006/covington.tracts10$P0010001),log(income.cov),pch=16,col="darkgrey",
     main="Covington Income Distribution vs.\n Percentage of Asian Population \n in 2010",
     xlab="Log Percentage of Asian Population",ylab="Log Median Income ($)")
abline(h=log(mean(income.cov,na.rm=TRUE)),col="firebrick",lwd=1.2)
abline(v=log(asian.ave.cov),col="darkgreen",lwd=1.2)
log.percent.as.cov = log(covington.tracts10$P0080007/covington.tracts10$P0010001)
log.percent.as.cov[which(log.percent.as.cov==-Inf)]=NA
abline(lm(log(income.cov)~log.percent.as.cov),col="navy",lwd=1.2)
text(-3,11.2, "Log Mean City Income", col = "firebrick",cex=0.8,font=2)
text(-4.9,10, "Log Average Percentage of Asian",col = "darkgreen", srt=90,cex=0.8,font=2)
legend("bottomleft","Linear Trend Between \n Income and Percentage of Asian",
       col="navy",lwd=1.2,cex=0.8,bty="n")


plot(log(covington.tracts10$P0070009/covington.tracts10$P0010001),log(income.cov),pch=16,col="darkgrey",
     main="Covington Income Distribution vs.\n Percentage of Hispanic Population \n in 2010",
     xlab="Log Percentage of Hispanic Population",ylab="Log Median Income ($)")
abline(h=log(mean(income.cov,na.rm=TRUE)),col="firebrick",lwd=1.2)
abline(v=log(hispanic.ave.cov),col="darkgreen",lwd=1.2)
log.percent.hi.cov = log(covington.tracts10$P0070009/covington.tracts10$P0010001)
log.percent.hi.cov[which(log.percent.hi.cov==-Inf)]=NA
abline(lm(log(income.cov)~log.percent.hi.cov),col="navy",lwd=1.2)
text(-6,11.2, "Log Mean City Income", col = "firebrick",cex=0.8,font=2)
text(-3.7,10, "Log Average Percentage of Hispanic",col = "darkgreen", srt=90,cex=0.8,font=2)
legend("bottomleft","Linear Trend Between \n Income and Percentage of Hispanic",
       col="navy",lwd=1.2,cex=0.8,bty="n")
dev.off()

#choropleth of Cincinnati and Covington combined, colored by median income level.
#We know from the race distribution plots above that Covington has a larger percentage
#of white population than Cincinnati; and we would expect some differences in income
#levels due to the differences in racial distribution.

#The orignial version as in website
totalColor = function(input){
  input[which(is.na(input))] = 0
  sapply(input, function(kk) rgb(1-kk/max(input), 1, 1-kk/max(input)))
}

png("Choropleth Cin & Cov Median Income.png",width=1500,height=1500,res=200)
```

```
col.cin <-totalColor(income.cin)
col.cov <-totalColor(income.cov)
plot(rbind(cincinnati.tracts10,covington.tracts10), col=c(col.cin,col.cov))
title("Ohio (Cincinnati) & Kentucky (Covington) \n colored by Median Income in 2010")

#Legend
i.min.cin = which(income.cin == min(income.cin))
i.med.cin = which(income.cin == median(income.cin))
i.max.cin = which(income.cin == max(income.cin))

i.min.cov = which(income.cov == min(income.cov))
i.med.cov = which(income.cov == median(income.cov))
i.max.cov = which(income.cov == max(income.cov))
legend("bottomright", paste("~", c(income.cin[c(i.min.cin, i.med.cin, i.max.cin)],
                        income.cov[c(i.min.cov, i.med.cov, i.max.cov)]), " Dollars", sep = ""),
      col = c(col.cin[c(i.min.cin, i.med.cin, i.max.cin)],col.cov[c(i.min.cov, i.med.cov, i.max.cov)]),pch = 16, lty =
0)

dev.off()

#Improved- now includes range and percentile

#top10%, top25%, above average, bottom 25%, bottom10%
income.share.cin<-na.omit(income.cin)/sum(na.omit(income.cin))*100
income.share.cov<-na.omit(income.cov)/sum(na.omit(income.cov))*100

#color from dark to light
col.income.cin<-ifelse(income.share.cin>0.9,"darkgreen",ifelse(income.share.cin>0.75,"palegreen4",
ifelse(income.share.cin>0.5,"palegreen3",ifelse(income.share.cin<0.25,"palegreen1","lightgreen"))))
col.income.cov<-ifelse(income.share.cov>0.9,"darkgreen",ifelse(income.share.cov>0.75,"palegreen4",
ifelse(income.share.cov>0.5,"palegreen3",ifelse(income.share.cov<0.25,"palegreen1","lightgreen"))))

col.cin<-col.income.cin[1:length(income.share.cin)]
col.cov<-col.income.cov[1:length(income.share.cov)]

png("Choropleth Cin & Cov Median Income.png",width=1500,height=1500,res=200)
plot(rbind(cincinnati.tracts10,covington.tracts10), col=c(col.cin,col.cov))
title("Ohio (Cincinnati) & Kentucky (Covington) \n colored by Median Income by Percenatge in 2010")
legend("bottomright", legend=c("Top 10%","Top 25%","Above Average","Bottom 25%", "Bottom 10%"),
col=c("darkgreen","palegreen4","palegreen3","palegreen1","lightgreen"),pch=16)
dev.off()
```

```
####################
#To analyze the age distribution, we use population pyramid (pyramid() function from
#package epicalc). For each tract in Cincinnati/Covington, the median male and female
#ages are treated as if they are age of an individual and ploted into a population pyramid.
#By visualizing how the two pyramids are different from each other, we can get information
#about the population-age structure (by gender) for the two cities.

#install.packages("epicalc")
library(epicalc)

#Creating a vector containing all median age values
age.cin = c(cincinnati.tracts10$age.male,cincinnati.tracts10$age.female)
age.cov = c(covington.tracts10$age.male,covington.tracts10$age.female)
#Create a gender vector to label each of the age values as they are being used
#by pyramid()
female.cin = rep("Female",length(cincinnati.tracts10$age.female))
female.cov = rep("Female",length(covington.tracts10$age.female))
male.cin = rep("Male",length(cincinnati.tracts10$age.male))
male.cov = rep("Male",length(covington.tracts10$age.male))
gender.cin=c(male.cin,female.cin)
gender.cov=c(male.cov,female.cov)

#Generate population pyramid
png("Population Pyramid Cincinnati.png")
pyramid(age=age.cin,sex=gender.cin,
    main="Population Pyramid of Median Age Distribution \n by Gender for Tracts Belonging to Cincinnati",
    col.gender=c("cornflowerblue","firebrick1"))
text(104.8,18.37,"  Median Age")
dev.off()

png("Population Pyramid Covington.png")
pyramid(age=age.cov,sex=gender.cov,
    main="Population Pyramid of Median Age Distribution \n by Gender for Tracts Belonging to Covington",
    col.gender=c("cornflowerblue","firebrick1"))
text(29,13.4,"Median Age")
dev.off()

##Improvements after presentaion##
#Beanplot as alternative to population pyramid for age distribution
png("Beanplot Age Distriution.png",height=2000,width=6000,res=350)
par(mfrow=c(1,4))

beanplot(cincinnati.tracts10$age.male,ylim=c(0,90))
abline(h=mean(cincinnati.tracts10$age.male),col="red",lwd=2)
```

```
text(0.98,40,"Mean Age for Cincinnati Males",font=2,col="red")
title(main="Distribution of Median Male \n  Ages for Cincinnati Tracts",
      ylab="Median Age")

beanplot(cincinnati.tracts10$age.female,ylim=c(0,90))
abline(h=mean(cincinnati.tracts10$age.female),col="red",lwd=2)
text(0.98,41.8,"Mean Age for Cincinnati Females",font=2,col="red")
title(main="Distribution of Median Female \n  Ages for Cincinnati Tracts",
      ylab="Median Age")

beanplot(covington.tracts10$age.male,ylim=c(0,90))
abline(h=mean(covington.tracts10$age.male),col="red",lwd=2)
text(0.98,40,"Mean Age for Covington Males",font=2,col="red")
title(main="Distribution of Median Male \n  Ages for Covington Tracts",
      ylab="Median Age")

beanplot(covington.tracts10$age.female,ylim=c(0,90))
abline(h=mean(covington.tracts10$age.female),col="red",lwd=2)
text(0.98,41.8,"Mean Age for Covington Females",font=2,col="red")
title(main="Distribution of Median Female \n  Ages for Covington Tracts",
      ylab="Median Age")

dev.off()

#How is income associated with age and gender in the two cities? LOWESS Smoothing
png("Lowess Income vs. Age.png",height=800,width=800,res=120)
age.income.male.cin = cbind(cincinnati.tracts10$age.male,cincinnati.tracts10$income.male)
age.income.male.cin = na.omit(age.income.male.cin)
plot(lowess(age.income.male.cin[,1],age.income.male.cin[,2]),type="l",col="blue",
     lwd=2,xlab="Tract-wide Median Age",ylab="Tract-wide Median Income (Dollars)",
     main="Income vs. Age by Gender and City")

age.income.female.cin = cbind(cincinnati.tracts10$age.female,cincinnati.tracts10$income.female)
age.income.female.cin = na.omit(age.income.female.cin)
lines(lowess(age.income.female.cin[,1],age.income.female.cin[,2]),col="red",lwd=2)

age.income.male.cov = cbind(covington.tracts10$age.male,covington.tracts10$income.male)
age.income.male.cov = na.omit(age.income.male.cov)
lines(lowess(age.income.male.cov[,1],age.income.male.cov[,2]),col="darkgreen",lwd=2)

age.income.female.cov = cbind(covington.tracts10$age.female,covington.tracts10$income.female)
age.income.female.cov = na.omit(age.income.female.cov)
lines(lowess(age.income.female.cov[,1],age.income.female.cov[,2]),col="darkorange2",lwd=2)
```

```
legend("bottomright",c("Cincinnati Males","Cincinnati Females","Covington Males","Covington Females"),
    col=c("blue", "red","darkgreen","darkorange2"),lwd=2,bty="n")

dev.off()
```