Untitled

2023-04-19

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

# Clearing data  
rm(list = ls())

Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.

# loading packages  
library(cancensus) #this package allows us to download census data from Stats Canada  
library(ggplot2)   
library(tidyverse) #collection of packages for data visualization and manipulation (includes dplyr and ggplot2)

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.0 ✔ readr 2.1.4  
## ✔ forcats 1.0.0 ✔ stringr 1.5.0  
## ✔ lubridate 1.9.2 ✔ tibble 3.2.0  
## ✔ purrr 1.0.1 ✔ tidyr 1.3.0  
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the ]8;;http://conflicted.r-lib.org/conflicted package]8;; to force all conflicts to become errors

library(tidycensus)  
library(tigris)

## To enable caching of data, set `options(tigris\_use\_cache = TRUE)`  
## in your R script or .Rprofile.

library(kableExtra)

## Warning in !is.null(rmarkdown::metadata$output) && rmarkdown::metadata$output  
## %in% : 'length(x) = 3 > 1' in coercion to 'logical(1)'

##   
## Attaching package: 'kableExtra'  
##   
## The following object is masked from 'package:dplyr':  
##   
## group\_rows

library(plotly)

##   
## Attaching package: 'plotly'  
##   
## The following object is masked from 'package:ggplot2':  
##   
## last\_plot  
##   
## The following object is masked from 'package:stats':  
##   
## filter  
##   
## The following object is masked from 'package:graphics':  
##   
## layout

library(sf)

## Linking to GEOS 3.10.2, GDAL 3.4.2, PROJ 8.2.1; sf\_use\_s2() is TRUE

library(spatialreg)

## Loading required package: spData  
## To access larger datasets in this package, install the spDataLarge  
## package with: `install.packages('spDataLarge',  
## repos='https://nowosad.github.io/drat/', type='source')`  
## Loading required package: Matrix  
##   
## Attaching package: 'Matrix'  
##   
## The following objects are masked from 'package:tidyr':  
##   
## expand, pack, unpack

library(spdep)

##   
## Attaching package: 'spdep'  
##   
## The following objects are masked from 'package:spatialreg':  
##   
## get.ClusterOption, get.coresOption, get.mcOption,  
## get.VerboseOption, get.ZeroPolicyOption, set.ClusterOption,  
## set.coresOption, set.mcOption, set.VerboseOption,  
## set.ZeroPolicyOption

library(spgwr)

## Loading required package: sp  
## NOTE: This package does not constitute approval of GWR  
## as a method of spatial analysis; see example(gwr)

library(dplyr)  
library(r5r)

## Please make sure you have already allocated some memory to Java by running:  
## options(java.parameters = '-Xmx2G').  
## You should replace '2G' by the amount of memory you'll require. Currently, Java memory is set to

library(rJava)  
library(cartogram)  
library(kableExtra)   
library(plotly) #packaged used to create interactive graphics   
library(sf) #package to work with geospatial data of simple feature class  
library(spatialreg) #package used to perform spatial regression analysis  
library(spdep) #package used to compute spatial dependence, weighting schemes, and spatial statistics  
library(spgwr) # package for geographically weighted regression   
library(dplyr) #package to work with and manipulate data frames  
library(r5r) #package to generate routing analysis and calculate the accessibility of parks in Vancouver  
library(rJava) #an r to java interface used to support the r5r package   
library(stargazer) #package to show results of regression analysis as a table

##   
## Please cite as:   
##   
## Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.  
## R package version 5.2.3. https://CRAN.R-project.org/package=stargazer

#Introduction

#Background

#Study Area This study took place at the Census Subdivision level to target the municipality of Vancouver, BC. There are 127 census tracts in Vancouver that were studied.

# storing the API key in system environment to be easily used in API calls  
options(cancensus.api\_key = "CensusMapper\_ff4c917e80df34b71900e0a048c999dd")

# assigning a persistent local cache  
options(cancensus.cache\_path = "/Users/dainadeangelis/Desktop/4GA3Project-/Project data")  
Sys.setenv(CM\_CACHE\_PATH = "/Users/dainadeangelis/Desktop/4GA3Project-/Project data")

# returning census data and geography for the Vancouver CMA as an sf-class dataframe  
census\_data <- get\_census(dataset='CA21', regions=list(CSD="5915022"), vectors=c("v\_CA21\_1","v\_CA21\_386","v\_CA21\_906", "v\_CA21\_4875", "v\_CA21\_4914", "v\_CA21\_1010"), labels="detailed", geo\_format="sf", level='CT')

## Reading vectors data from local cache.

## Reading geo data from local cache.

# removing columns that are not needed for analysis  
census\_data <- census\_data[,-c(4,5,9,10,13,15,16)]

# renaming the columns in the dataset  
names(census\_data) [1] <- "Shape\_area\_km2"  
names(census\_data) [3] <- "Number\_of\_households"  
names(census\_data) [4] <- "Region\_name"  
names(census\_data) [7] <- "Number\_of\_dwellings"  
names(census\_data) [10:15] <- c("Population\_2021", "Average\_age", "Median\_total\_income\_of\_household\_2020", "Total\_visible\_minority\_population", "Not\_a\_visible\_minority", "Low\_income\_status\_for\_the\_population\_in\_private\_households\_2020")

#Data The data used for analysis includes census data for variables of population, population density, visible minority status, non-visible minority status, and income for each census tract in the census subdivision of Vancouver. This data was obtained from Statistics Canada through use of the cancensus package in R. All census data used is from the 2021 census. Point data for each the location of each park, homeless shelter, and street tree in the census subdivision was downloaded in a .csv file format from the City of Vancouver’s open data portal. To use the r5r package so that accessibility to parks by both walking and public transit could be determined, a road network dataset of Vancouver, and as well as a public transport feed of the city was needed. The road network dataset was obtained from BBBike, and was stored as a .pbf file. The public transport feed was obtained from Transitland and stored in a GTFS.zip file.

#Methods To determine the relationship between greenspace and other socioeconomic factors in Vancouver, Rstudio was used to analyze and visualize the variables being examined. Choropleth maps for variables of population, population density, visible minority status, non-visible minority status, and income, and accessibility to parks were made to provide a visual representation of the difference between census tracts. Regression analysis was used to determine the relationship between the independent variable studied (number of parks accessible within a 30 minute travel time) and dependent variables (population, population density, visible minority status, non-visible minority status, and income). Scatterplots were used to show a visual representation between independent and dependent variables studied to better understand the results of regression analysis.

#Results

To begin our analysis, the centroids for each census tract in the Vancouver CSD were obtained and put into a dataframe with latitude, longitude, and census ID. This allowed us to have point data describing the middle of each CT.

# obtaining centroids for census tracts in the Vancouver CSD   
CT\_centroids <- st\_coordinates(st\_centroid(census\_data))  
#converting centroids to data frame  
centroids\_df <- as.data.frame(CT\_centroids)  
class(centroids\_df)

## [1] "data.frame"

centroids\_df<- cbind(centroids\_df, census\_data$GeoUID)  
  
centroids\_df <- centroids\_df %>%   
 rename("lon" = "X",  
 "lat" = "Y",  
 "id" = "census\_data$GeoUID")

# calculating population density for the census tracts in Vancouver  
census\_data <- census\_data %>%  
 mutate(Population\_density =  
 (Population\_2021 / Shape\_area\_km2))

# calculating the proportion of the minority population and adding to census\_data  
census\_data <- census\_data %>%  
 mutate(Proportion\_visible\_minority= (Total\_visible\_minority\_population / Population\_2021))

# calculating the proportion of the non visible minority population and adding to census\_data  
census\_data <- census\_data %>%  
 mutate(Proportion\_nonvisible\_minority= (Not\_a\_visible\_minority / Population\_2021))

# calculating the proportion of low income status in Vancouver census tracts  
census\_data <- census\_data %>%  
 mutate(Proportion\_low\_income = (Low\_income\_status\_for\_the\_population\_in\_private\_households\_2020 / Population\_2021))

The point data for park in the Vancouver CSD was obtained from a csv file from Vancouver’s open data portal and transformed into a data frame with columns describing the X and Y coordinates of each park.

# reading csv files  
parks <- read.csv(file = "Project data/parks.csv", head = TRUE, sep = ";")

#extracting point data from park csv file  
parks\_coord <- data.frame(do.call("rbind", strsplit(as.character(parks$GoogleMapDest), ",", fixed = TRUE)))  
  
# Rename multiple columns  
parks\_coord <- parks\_coord %>%   
 rename("X" = "X1",  
 "Y" = "X2")

The point data for parks was prepared for use of the r5r package by renaming coulumns for latitude, longitude, and park ID data and to the names required by r5r. Data for latitude and longitude was converted into a numeric format. Lastly, a column to indicate that each park point was one singular park was added to meet the requirement of the opportunites parameter in r5r.

#preparing park data for r5r  
class(parks\_coord)

## [1] "data.frame"

parks\_coord<- cbind(parks\_coord, parks$ParkID)  
  
# renaming columns  
parks\_coord <- parks\_coord %>%   
 rename("lat" = "X",  
 "lon" = "Y",  
 "id" = "parks$ParkID")  
  
# changing columns to numeric  
parks\_coord$lon <- as.numeric(parks\_coord$lon)  
parks\_coord$lat <- as.numeric(parks\_coord$lat)

parks\_coord <- parks\_coord %>%  
 add\_column(Parks\_count = 1)

Next, the r5r package was used in order to build a transport network for Vancouver, so that the distances and routing between the centroids of census tracts and park locations can be calculated.

#allocating RAM memory to Java to use the r5r package  
options(java.parameters = "-Xmx2G")

The transport network was built with using a road network dataset of Vancouver, and as well as a public transport feed of the city. The road network dataset was obtained from BBBike, and was stored as a .pbf file. The public transport feed was obtained from Transitland and stored in a GTFS.zip file.

#building transport network for r5r  
path <- file.path("Project data/r5rpath")  
  
#building transport network for r5r  
list.files(path)

## [1] "GTFS\_TRANIST.zip" "network\_settings.json"   
## [3] "network.dat" "Vancouver.osm.pbf"   
## [5] "Vancouver.osm.pbf.mapdb" "Vancouver.osm.pbf.mapdb.p"

r5r\_core <- setup\_r5(data\_path = path)

## No raster .tif files found. Using elevation = 'NONE'.

## Using cached R5 version from /Library/Frameworks/R.framework/Versions/4.2/Resources/library/r5r/jar/r5-v6.9-all.jar

##   
## Using cached network.dat from Project data/r5rpath/network.dat

#load origin points as the centroids of census tracts  
points <- centroids\_df  
class(points)

## [1] "data.frame"

#points of interest: parks  
poi <- parks\_coord  
#departure times  
departure\_datetime <- as.POSIXct("16-02-2023 14:00:00",  
 format = "%d-%m-%Y %H:%M:%S")

The accessibility function in r5r was used to compute how many parks were accessible within 30 minutes of each census tracts centroid by walking or public transit.

#using accessibility from r5r  
access <- accessibility(r5r\_core = r5r\_core,  
 origins = points,  
 destinations = poi,  
 opportunities\_colnames = c("Parks\_count"),  
 mode = c("WALK", "TRANSIT"),  
 departure\_datetime = departure\_datetime,  
 decay\_function = "step",  
 cutoffs = 30)

## Warning in assign\_points\_input(destinations, "destinations"): 'destinations$id'  
## forcefully cast to character.

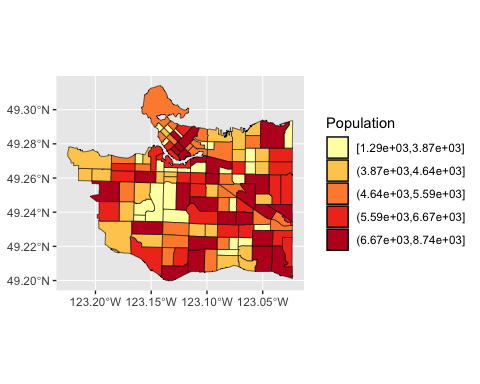
Once the number of accessible parks was calculated, this was added to the data frame displaying census data.

census\_data <- cbind(census\_data, accessibility = access$accessibility)

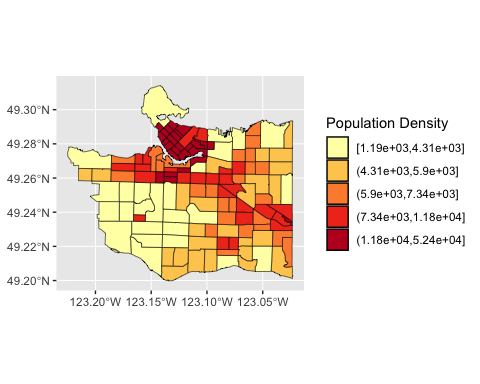
After all variables were in one data frame and prepared for analysis, data was first visualized by creating choropleth maps for each variable of interest. Maps were created for each variable as follows:

Population of each CT

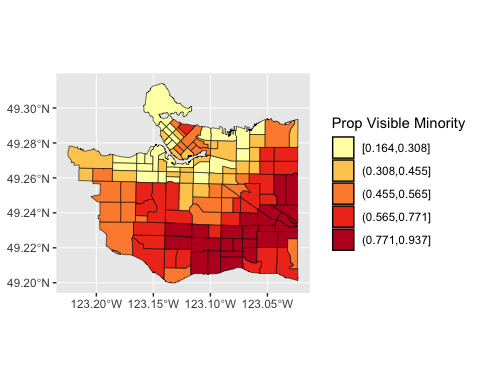
# creating a choropleth map for the 2021 population  
ggplot(census\_data) +  
 geom\_sf(aes(fill = cut\_number(Population\_2021, 5)),  
 colour = "black",  
 size = 0.1) +  
 scale\_fill\_brewer(palette = "YlOrRd") +  
 coord\_sf() +  
 labs(fill = "Population")

 Population density of each CT

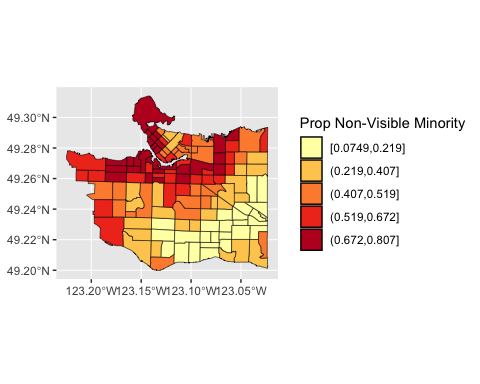
# creating a choropleth map for the population density  
ggplot(census\_data) +  
 geom\_sf(aes(fill = cut\_number(Population\_density, 5)),  
 colour = "black",  
 size = 0.1) +  
 scale\_fill\_brewer(palette = "YlOrRd") +  
 coord\_sf() +  
 labs(fill = "Population Density")

 Proportion of population identifying as visible and non-visible minorities

# creating a choropleth map for the proportion of the visisble minority population  
ggplot(census\_data) +  
geom\_sf(aes(fill = cut\_number(Proportion\_visible\_minority, 5)),  
 color = "black",  
 size = 0.1) +  
scale\_fill\_brewer(palette = "YlOrRd") +  
labs(fill = "Prop Visible Minority")

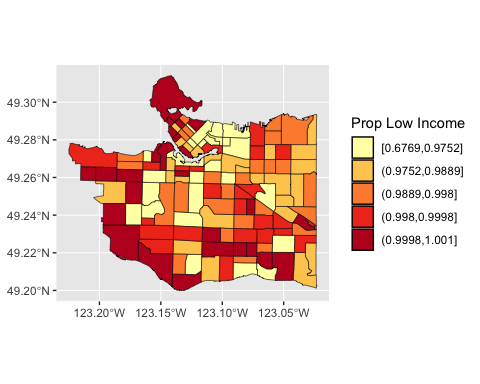


# creating a choropleth map for the proportion of the non-visible minority population  
ggplot(census\_data) +  
geom\_sf(aes(fill = cut\_number(Proportion\_nonvisible\_minority, 5)),  
 color = "black",  
 size = 0.1) +  
scale\_fill\_brewer(palette = "YlOrRd") +  
labs(fill = "Prop Non-Visible Minority")



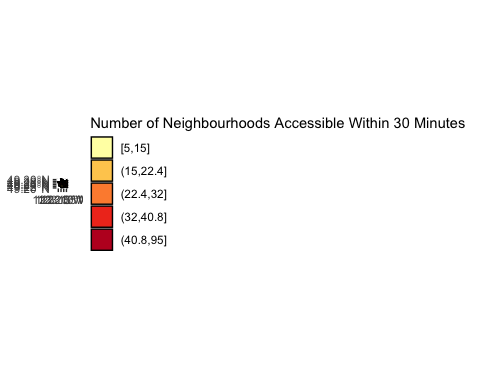
The proportion of the population considered low income for each CT

# creating a choropleth map displaying the proportion of low-income   
ggplot(census\_data) +  
geom\_sf(aes(fill = cut\_number(Proportion\_low\_income, 5)),  
 color = "black",  
 size = 0.1) +  
scale\_fill\_brewer(palette = "YlOrRd") +  
labs(fill = "Prop Low Income")



Finally, a choropleth map was created to display the independent variable, parks accessible within 30 mintues of a centroid.

# creating a choropleth map displaying the number of parks accessible to a census tract's centroids within 30 minutes  
ggplot(census\_data) +  
geom\_sf(aes(fill = cut\_number(accessibility, 5)),  
 color = "black",  
 size = 0.1) +  
scale\_fill\_brewer(palette = "YlOrRd") +  
labs(fill = "Number of Neighbourhoods Accessible Within 30 Minutes")



Next, to determine the relationship between the number of accessible parks and each independent variable, the independent variables were regressed to number of accessible parks for each CT.

#creating a regression model of population regressing on number of accessible parks  
model\_population <- lm(formula = Population\_2021 ~ accessibility,   
 data = census\_data)  
  
stargazer(model\_population,  
 header = FALSE,  
 title = "Population of Census Tracts regressed on Number of Parks Accessible")

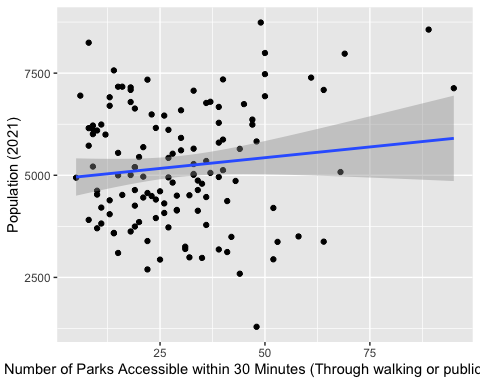
##   
## \begin{table}[!htbp] \centering   
## \caption{Population of Census Tracts regressed on Number of Parks Accessible}   
## \label{}   
## \begin{tabular}{@{\extracolsep{5pt}}lc}   
## \\[-1.8ex]\hline   
## \hline \\[-1.8ex]   
## & \multicolumn{1}{c}{\textit{Dependent variable:}} \\   
## \cline{2-2}   
## \\[-1.8ex] & Population\\_2021 \\   
## \hline \\[-1.8ex]   
## accessibility & 10.518 \\   
## & (7.816) \\   
## & \\   
## Constant & 4,904.639$^{\*\*\*}$ \\   
## & (264.208) \\   
## & \\   
## \hline \\[-1.8ex]   
## Observations & 127 \\   
## R$^{2}$ & 0.014 \\   
## Adjusted R$^{2}$ & 0.006 \\   
## Residual Std. Error & 1,459.451 (df = 125) \\   
## F Statistic & 1.811 (df = 1; 125) \\   
## \hline   
## \hline \\[-1.8ex]   
## \textit{Note:} & \multicolumn{1}{r}{$^{\*}$p$<$0.1; $^{\*\*}$p$<$0.05; $^{\*\*\*}$p$<$0.01} \\   
## \end{tabular}   
## \end{table}

summary(model\_population)

##   
## Call:  
## lm(formula = Population\_2021 ~ accessibility, data = census\_data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -4118.5 -1065.2 -159.5 1192.0 3319.0   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4904.639 264.208 18.564 <2e-16 \*\*\*  
## accessibility 10.518 7.816 1.346 0.181   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1459 on 125 degrees of freedom  
## Multiple R-squared: 0.01428, Adjusted R-squared: 0.006395   
## F-statistic: 1.811 on 1 and 125 DF, p-value: 0.1808

For each regression, a scatter plot was created to provide a visual representation of the data

ggplot(data = census\_data,   
 aes(x = accessibility,   
 y = Population\_2021))+  
 geom\_point() +  
 geom\_smooth(formula = y ~ x,  
 method = "lm") +  
 ylab("Population (2021)") +  
 xlab("Number of Parks Accessible within 30 Minutes (Through walking or public transit")



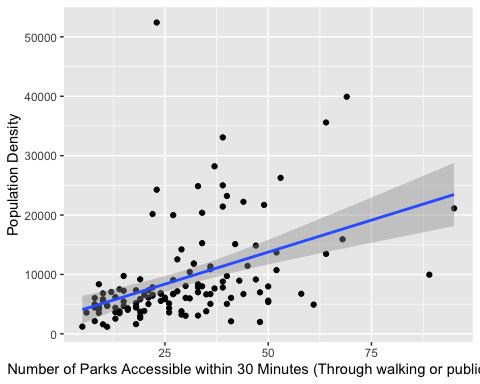
Model regressing population density on number of accessible marks

#creating a regression model of population density regressing on number of accessible parks  
model\_population\_density <- lm(formula = Population\_density ~ accessibility,   
 data = census\_data)  
  
stargazer(model\_population\_density,  
 header = FALSE,  
 title = "Population Density of Census Tracts regressed on Number of Parks Accessible")

##   
## \begin{table}[!htbp] \centering   
## \caption{Population Density of Census Tracts regressed on Number of Parks Accessible}   
## \label{}   
## \begin{tabular}{@{\extracolsep{5pt}}lc}   
## \\[-1.8ex]\hline   
## \hline \\[-1.8ex]   
## & \multicolumn{1}{c}{\textit{Dependent variable:}} \\   
## \cline{2-2}   
## \\[-1.8ex] & Population\\_density \\   
## \hline \\[-1.8ex]   
## accessibility & 215.382$^{\*\*\*}$ \\   
## & (39.775) \\   
## & \\   
## Constant & 2,995.220$^{\*\*}$ \\   
## & (1,344.555) \\   
## & \\   
## \hline \\[-1.8ex]   
## Observations & 127 \\   
## R$^{2}$ & 0.190 \\   
## Adjusted R$^{2}$ & 0.184 \\   
## Residual Std. Error & 7,427.142 (df = 125) \\   
## F Statistic & 29.322$^{\*\*\*}$ (df = 1; 125) \\   
## \hline   
## \hline \\[-1.8ex]   
## \textit{Note:} & \multicolumn{1}{r}{$^{\*}$p$<$0.1; $^{\*\*}$p$<$0.05; $^{\*\*\*}$p$<$0.01} \\   
## \end{tabular}   
## \end{table}

Scatterplot of population density vs number of accessible parks

ggplot(data = census\_data,   
 aes(x = accessibility,   
 y = Population\_density))+  
 geom\_point() +  
 geom\_smooth(formula = y ~ x,  
 method = "lm") +  
 ylab("Population Density") +  
 xlab("Number of Parks Accessible within 30 Minutes (Through walking or public transit")



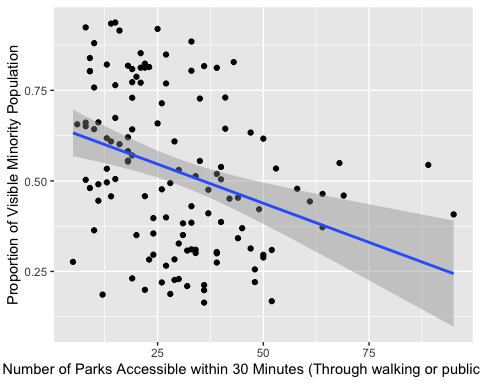
Model regressing proportion of visible minorities on number of accessible parks

#creating a regression model of population density regressing on number of accessible parks  
model\_visible\_minority <- lm(formula = Proportion\_visible\_minority ~ accessibility,   
 data = census\_data)  
  
stargazer(model\_visible\_minority,  
 header = FALSE,  
 title = "Proportion of Visible Minority Population in Census Tracts regressed on Number of Parks Accessible")

##   
## \begin{table}[!htbp] \centering   
## \caption{Proportion of Visible Minority Population in Census Tracts regressed on Number of Parks Accessible}   
## \label{}   
## \begin{tabular}{@{\extracolsep{5pt}}lc}   
## \\[-1.8ex]\hline   
## \hline \\[-1.8ex]   
## & \multicolumn{1}{c}{\textit{Dependent variable:}} \\   
## \cline{2-2}   
## \\[-1.8ex] & Proportion\\_visible\\_minority \\   
## \hline \\[-1.8ex]   
## accessibility & $-$0.004$^{\*\*\*}$ \\   
## & (0.001) \\   
## & \\   
## Constant & 0.654$^{\*\*\*}$ \\   
## & (0.037) \\   
## & \\   
## \hline \\[-1.8ex]   
## Observations & 127 \\   
## R$^{2}$ & 0.110 \\   
## Adjusted R$^{2}$ & 0.103 \\   
## Residual Std. Error & 0.205 (df = 125) \\   
## F Statistic & 15.511$^{\*\*\*}$ (df = 1; 125) \\   
## \hline   
## \hline \\[-1.8ex]   
## \textit{Note:} & \multicolumn{1}{r}{$^{\*}$p$<$0.1; $^{\*\*}$p$<$0.05; $^{\*\*\*}$p$<$0.01} \\   
## \end{tabular}   
## \end{table}

Scatterplot of proportion of visible minorities vs number of accessible parks

ggplot(data = census\_data,   
 aes(x = accessibility,   
 y = Proportion\_visible\_minority))+  
 geom\_point() +  
 geom\_smooth(formula = y ~ x,  
 method = "lm") +  
 ylab("Proportion of Visible Minority Population") +  
 xlab("Number of Parks Accessible within 30 Minutes (Through walking or public transit")



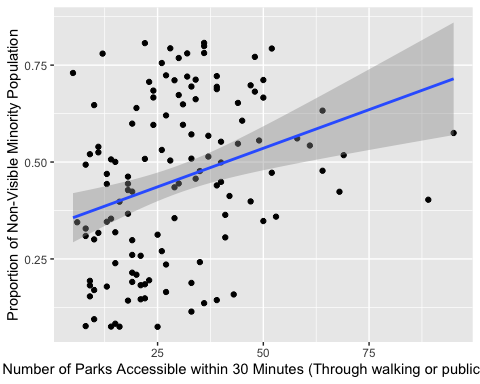
Model regressing proportion of non-visible minorities on number of accessible parks

#creating a regression model of population density regressing on number of accessible parks  
model\_nonvisible\_minority <- lm(formula = Proportion\_nonvisible\_minority ~ accessibility,   
 data = census\_data)  
  
stargazer(model\_nonvisible\_minority,  
 header = FALSE,  
 title = "Proportion of Non-Visible Minority Population in Census Tracts regressed on Number of Parks Accessible")

##   
## \begin{table}[!htbp] \centering   
## \caption{Proportion of Non-Visible Minority Population in Census Tracts regressed on Number of Parks Accessible}   
## \label{}   
## \begin{tabular}{@{\extracolsep{5pt}}lc}   
## \\[-1.8ex]\hline   
## \hline \\[-1.8ex]   
## & \multicolumn{1}{c}{\textit{Dependent variable:}} \\   
## \cline{2-2}   
## \\[-1.8ex] & Proportion\\_nonvisible\\_minority \\   
## \hline \\[-1.8ex]   
## accessibility & 0.004$^{\*\*\*}$ \\   
## & (0.001) \\   
## & \\   
## Constant & 0.336$^{\*\*\*}$ \\   
## & (0.037) \\   
## & \\   
## \hline \\[-1.8ex]   
## Observations & 127 \\   
## R$^{2}$ & 0.097 \\   
## Adjusted R$^{2}$ & 0.090 \\   
## Residual Std. Error & 0.203 (df = 125) \\   
## F Statistic & 13.473$^{\*\*\*}$ (df = 1; 125) \\   
## \hline   
## \hline \\[-1.8ex]   
## \textit{Note:} & \multicolumn{1}{r}{$^{\*}$p$<$0.1; $^{\*\*}$p$<$0.05; $^{\*\*\*}$p$<$0.01} \\   
## \end{tabular}   
## \end{table}

Scatterplot of proportion of non-visible minorities vs number of accessible parks

ggplot(data = census\_data,   
 aes(x = accessibility,   
 y = Proportion\_nonvisible\_minority))+  
 geom\_point() +  
 geom\_smooth(formula = y ~ x,  
 method = "lm") +  
 ylab("Proportion of Non-Visible Minority Population") +  
 xlab("Number of Parks Accessible within 30 Minutes (Through walking or public transit")



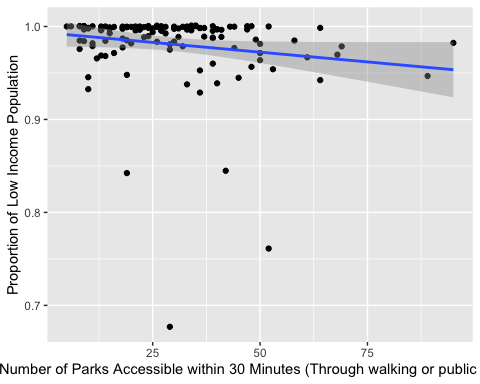
Model regressing proportion of low income population on number of accessible parks

#creating a regression model of population density regressing on number of accessible parks  
model\_income <- lm(formula = Proportion\_low\_income ~ accessibility,   
 data = census\_data)  
  
stargazer(model\_income,  
 header = FALSE,  
 title = "Proportion of Low Income Population in Census Tracts regressed on Number of Parks Accessible")

##   
## \begin{table}[!htbp] \centering   
## \caption{Proportion of Low Income Population in Census Tracts regressed on Number of Parks Accessible}   
## \label{}   
## \begin{tabular}{@{\extracolsep{5pt}}lc}   
## \\[-1.8ex]\hline   
## \hline \\[-1.8ex]   
## & \multicolumn{1}{c}{\textit{Dependent variable:}} \\   
## \cline{2-2}   
## \\[-1.8ex] & Proportion\\_low\\_income \\   
## \hline \\[-1.8ex]   
## accessibility & $-$0.0004$^{\*}$ \\   
## & (0.0002) \\   
## & \\   
## Constant & 0.993$^{\*\*\*}$ \\   
## & (0.008) \\   
## & \\   
## \hline \\[-1.8ex]   
## Observations & 127 \\   
## R$^{2}$ & 0.028 \\   
## Adjusted R$^{2}$ & 0.020 \\   
## Residual Std. Error & 0.041 (df = 125) \\   
## F Statistic & 3.571$^{\*}$ (df = 1; 125) \\   
## \hline   
## \hline \\[-1.8ex]   
## \textit{Note:} & \multicolumn{1}{r}{$^{\*}$p$<$0.1; $^{\*\*}$p$<$0.05; $^{\*\*\*}$p$<$0.01} \\   
## \end{tabular}   
## \end{table}

Scatterplot of proportion of low income population vs number of accessible parks

ggplot(data = census\_data,   
 aes(x = accessibility,   
 y = Proportion\_low\_income))+  
 geom\_point() +  
 geom\_smooth(formula = y ~ x,  
 method = "lm") +  
 ylab("Proportion of Low Income Population") +  
 xlab("Number of Parks Accessible within 30 Minutes (Through walking or public transit")



#Analysis

#Conclusion