Final Project: Walkability and Public Health in the US

Lou Godmer, Kienan Battin, Divakar Mehta

April 17, 2023

Contents

Objective

The objective is to quantify the causal effect that the "walkability" of a region has on public health. The original data comes from two sources: 1. The U.S. Chronic Disease Indicators provides reported cases of a set of 124 indicators that are important to public health, and the geographic location of the case. 2. The Walkability Index quantifies every Census 2019 block group's relative "Walkability" as defined by the EPA based on characteristics such as easy walking access to public transit, jobs, stores and services. Quantifying the causal effect of walkability on public health can help policy makers understand how community planning measures that may improve or degrade the walkability of the region will impact public health.

The appendix of this document describes the pre-processing methodology that was used combine the two data sets to enable the quantitative analysis. Because the pre-processing methodology can take an hour or more to execute, the pre-processed data was exported. The beginning of this document imports the pre-processed data and the rest of the analysis is done based on the pre-processed data.

Load necessariy libararies

```
rm(list=ls())

options(repos = list(CRAN="http://cran.rstudio.com/"))

if (!require('NHANES')) install.packages('NHANES')
library('openxlsx')

if (!require('ggplot2')) install.packages('ggplot2')
library('ggplot2')

if (!require('dplyr')) install.packages('dplyr')
library('dplyr')

if (!require('GGally')) install.packages('GGally')
library('GGally')

if (!require('tableone')) install.packages('tableone')
library(tableone)

if (!require('pROC')) install.packages('pROC')
library(pROC)

if (!require('tidycensus')) install.packages('tidycensus')
library(tidycensus)

if (!require('tigris')) install.packages('tigris')
library(tigris)

if (!require('sf')) install.packages('sf')
library(sf)

if (!require('stringr')) install.packages('stringr')
library(stringr)

if (!require('dplyr')) install.packages('dplyr')
library(dplyr)
```

Load the data

Download the data which has already undergone the pre-processing methodology described in the appendix. WARNING: this may take several minutes. To avoid unnecessary downloads, the commands are commented out. Un-comment and execute the commands to download the data.

```
#download.file("https://walkabilityandhealth.blob.core.windows.net/walkabilityandhealth/disease_with_wa
#unzip("disease_with_walkability.zip", "disease_with_walkability.csv")
```

```
disease_with_walkability <- read.csv("disease_with_walkability.csv")
```

Do some cleanup and data shaping

```
disease_with_walkability = filter(disease_with_walkability, !is.na(NatWalkInd))
nwi25 <- 5.83 # the bottom 25 percent - least walkable blocks in the US have index less than 5.83
nwi75 <- 13.17 # the top 25 percent- most walkable blocks in the US have index greater than 13.17
# analysis will only use the disease data from the least walkable and most walkable blocks
disease_with_walkability = filter(disease_with_walkability, NatWalkInd < nwi25 | NatWalkInd > nwi75)
disease_with_walkability$Walkable <- ifelse(disease_with_walkability$NatWalkInd >= nwi75, 0, 1)
```

```
disease_with_walkability$Gender = ifelse(disease_with_walkability$StratificationCategory1 == "Gender", 
disease_with_walkability$Race = ifelse(disease_with_walkability$StratificationCategory1 == "Race/Ethnic
```

Get familiar with the data

Descriptions of the fields in the dataset

The table below describes the fields that are used in this analysis

Source Data Set	Field Name	Field Description	Usage In This Analysis
Walkability	Walkable	Binary variable, 1 if the block is walkable, 0 otherwise	Treatment Variable
Disease Indicators	LocationAbbr	US State or Territory Abbreviation	
Disease Indicators	LocationDesc	US State or Territory name	
Disease Indicators	DataSource	Origin of the disease indicator data	
Disease Indicators	Topic	Category of the disease information, i.e. “Asthma”	
Disease Indicators	Question	Brief description of the condition being measured	Dependent variable category
Disease Indicators	DataValueUnit	Unit of measurement for the response to “Question”, i.e. “gallons”	

Source Data Set	Field Name	Field Description	Usage In This Analysis
Disease Indicators	DataValueType	Type of measurement for the response to "Question", i.e. "mean"	
Disease Indicators	DataValueAlt	Numeric value of the response to "Question"	Dependent variable value
Disease Indicators	Gender	Gender of the subject of the disease indicator data	Independent var (possible confounder)
Disease Indicators	Race	Race/Ethnicity of the subject of the disease indicator data	Independent var (possible confounder)
Disease Indicators	GeoLocation	Longitude and latitude of the location where the data was collected	
Disease Indicators	STATEFP	FIPS state code of the state of GeoLocation	
Disease Indicators	COUNTYFP	FIPS county code of the county of GeoLocation	
Disease Indicators	TRACTCE	FIPS tract code of the tract of GeoLocation	
Disease Indicators	BLKGRPCE	FIPS block code of the block group of GeoLocation	
Disease Indicators	GEOID	Full GEOID (state, county, tract, block group) of GeoLocation	
Walkability	CSA	"Combined Statistical Area" - grouping of adjacent metropolitan statistical areas that share social and economic ties	
Walkability	CSA_NAME	Friendly name of the CSA	
Walkability	CBSA	"Core Based Statistical Area" - functional region based around an urban center along with adjacent areas that are socioeconomically tied to the urban center by commuting	
Walkability	CBSA_NAME	Friendly name of the CBSA	
Walkability	CBSA_POP	Estimated population of the CBSA	
Walkability	CBSA_EMP	Total number of employees in the CBSA	
Walkability	CBSA_WRK	Total number of workers in the CBSA	

Source Data Set	Field Name	Field Description	Usage In This Analysis
Walkability	AC_Total	Total area of land in square meters within the block group	
Walkability	AC_Water	Total area of land in square meters covered by water within the block group	
Walkability	AC_Land	Total area of land in square meters not covered by water within the block group	
Walkability	AC_Unpr	Total are of land in square meters classified as unproductive or unused within the block group	
Walkability	TotPop	Total population within the block group	
Walkability	CountHU	Count of housing units in the block group	
Walkability	HH	Count of occupied housing units in the block group	
Walkability	P_WrkAge	Percentage of the population that is of working age (16 or older)	
Walkability	AutoOwn0	Households with zero automobiles	
Walkability	Pct_AO0	Percentage of households with zero automobiles	
Walkability	AutoOwn1	Households with one automobiles	
Walkability	Pct_AO1	Percentage of households with one automobiles	
Walkability	AutoOwn2p	Households with two or more automobiles	
Walkability	Pct_AO2p	Percentage of households with two or more automobiles	
Walkability	Workers	Population of workers (16 or older) in the block group	
Walkability	R_LowWageWk	Number of workers earning $1250/month or less (home location)	

Source Data Set	Field Name	Field Description	Usage In This Analysis
Walkability	R_MedWageWk	Number of workers earning more than $1250/month and less than $3333/month (home location)	
Walkability	R_HiWageWk	Number of workers earning $3333/month or more (home location)	
Walkability	R_PCTLOWWAGE	Low wage workers as a percent of all workers in CBG (home location)	
Walkability	TotEmp	Total employment	
Walkability	E8_Ret	Retail jobs within a 8-tier employment classification scheme	
Walkability	E8_off	Office jobs within a 8-tier employment classification scheme	
Walkability	E8_Ind	Industrial jobs within a 8-tier employment classification scheme	
Walkability	E8_Svc	Service jobs within a 8-tier employment classification scheme	
Walkability	E8_Ent	Entertainment jobs within a 8-tier employment classification scheme	
Walkability	E8_Ed	Education jobs within a 8-tier employment classification scheme	
Walkability	E8_Hlth Healthcare jobs within a 8-tier employment classification scheme		
Walkability	E8_Pub	Public administration jobs within a 8-tier employment classification scheme	
Walkability	E_LowWageWk	Number of workers earning $1250/month or less (work location)	
Walkability	E_MedWageWk	Number of workers earning more than $1250/month and less than $3333/month (work location)	
Walkability	E_HiWageWk	Number of workers earning $3333/month or more (work location)	

Source Data Set	Field Name	Field Description	Usage In This Analysis
Walkability	E_PctLowWage	Low wage workers as a percent of all workers in CBG (work location)	
Walkability	D1A	Gross residential density (HU/acre) on unprotected land	
Walkability	D1B	Gross population density (people/acre) on unprocted land	
Walkability	D1C	Gross employment density (jobs/acre) on unprotected land	
Walkability	D1C8_RET	Gross retail (8-tier) employment density (jobs/acre) on unprotected land	
Walkability	D1C8_OFF	Gross office (8-tier) employment density (jobs/acre) on unprotected land	
Walkability	D1C8_IND	Gross industrial (8-tier) employment density (jobs/acre) on unprotected land	
Walkability	D1C8_SVC	Gross service (8-tier) employment density (jobs/acre) on unprotected land	
Walkability	D1C8_ENT	Gross entertainment (8-tier) employment density (jobs/acre) on unprotected land	
Walkability	D1C8_ED	Gross education (8-tier) employment density (jobs/acre) on unprotected land	
Walkability	D1C8_HLTH	Gross healthcare (8-tier) employment density (jobs/acre) on unprotected land	
Walkability	D1C8_PUB	Gross public administration (8-tier) employment density (jobs/acre) on unprotected land	
Walkability	D1D	Gross activity density (HU + employment / acre) on unprotected land	
Walkability	D2A_JPHH	Jobs per housing unit	
Walkability	D2B_E8MIX	8-tier employment entropy	

Source Data Set	Field Name	Field Description	Usage In This Analysis
Walkability	D2B_E8MIXA	8-tier employment entropy, denominator set to the static 8 employment types in the CBG	
Walkability	D2C_TRPMX2	Employment and household entropy (excluding industrial jobs), based on trip production and attraction	
Walkability	D2C_TRIPEQ	Trip production and trip attractions equilibrium index (closer to 1 = more balance)	
Walkability	D2R_JOBPOP	Deviation of CBG jobs/population ratio from regional average jobs/pop ratio	
Walkability	D2R_WRKEMP	Household workers per job	
Walkability	D2A_WORKEMP	Deviation of CBG ratio of household workers/job from regional average ratio of household workers/ob	
Walkability	D2C_WREMLX	Household worker per job equilibrium index (closer to one = more balanced)	
Walkability	D4A	Distance from population weighted centroid to nearest transit stop, meters	
Walkability	D4B025	Proportion of CBG employment within 1/4 mile of fixed guideway transit stop	
Walkability	D4B050	Proportion of CBG employment within 1/2 mile of fixed guideway transit stop	
Walkability	D4C	Transit service frequency. (Afternoon peak period transit departure within 0.25 miles)	
Walkability	D4D	Peak pm transit departure within 0.25 miles of CBG, per square mile	

Source Data Set	Field Name	Field Description	Usage In This Analysis
Walkability	D5AR	Jobs within a 45 minute drive (weighted)	
Walkability	D5AE	Working-age population within 45 min. drive (weighted)	
Walkability	D5BR	Jobs within 45 min. transit commute (weighted)	
Walkability	D5BE	Working-age population within 45 min. transit commute (weighted)	
Walkability	D5CR	Job accessibility (D5ar) as proportion of total regional job accessibility	
Walkability	D5CRI	Regional centrality index (auto) - D5cr divided by max D5cr in metro region (CBSA)	
Walkability	D5CE	Accessibility to working-age populatin (D5ae) as proportion of total regional accessibility	
Walkability	D5CEI	Regional centrality index (auto) - D5ce divided by max D5ce in metro region (CBSA)	
Walkability	D5DR	Job accessibility by transit (D5br) as proportion of total regional job accessibility by transit	
Walkability	D5DRI	Regional centrality index (transit) - D5dr divided by max D5dr in metro region (CBSA)	
Walkability	D5DE	Accessibility to working-age populatin by transit (D5be) as proportion of total regional accessibility	
Walkability	D5DEI	Regional centrality index (transit) - D5de divided by max D5de in metro region (CBSA)	

```
str(disease_with_walkability)
```

```
## 'data.frame':    300936 obs. of  158 variables:
##  $ X                         : int  1 7 15 20 26 39 40 42 51 56 ...
##  $ YearStart                 : int  2014 2013 2020 2017 2020 2013 2014 2008 2012 2008 ...
```

```
##  $ YearEnd                  : int  2014 2013 2020 2017 2020 2017 2018 2012 2016 2012 ...
##  $ LocationAbbr             : chr  "AR" "OR" "IL" "MA" ...
##  $ LocationDesc             : chr  "Arkansas" "Oregon" "Illinois" "Massachusetts" ...
##  $ DataSource               : chr  "SEDD; SID" "SEDD; SID" "NVSS" "NVSS" ...
##  $ Topic                    : chr  "Asthma" "Asthma" "Asthma" "Asthma" ...
##  $ Question                 : chr  "Hospitalizations for asthma" "Hospitalizations for asthma" "Asthm
##  $ Response                 : logi  NA NA NA NA NA NA ...
##  $ DataValueUnit            : chr  "" "" "Number" "" ...
##  $ DataValueType            : chr  "Number" "Number" "Number" "Number" ...
##  $ DataValue                : chr  "916" "760" "89" "28" ...
##  $ DataValueAlt             : num  916 760 89 28 26 ...
##  $ DataValueFootnoteSymbol  : chr  "" "" "" "" ...
##  $ DatavalueFootnote        : chr  "" "" "" "" ...
##  $ LowConfidenceLimit       : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ HighConfidenceLimit      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ StratificationCategory1  : chr  "Gender" "Gender" "Gender" "Gender" ...
##  $ Stratification1          : chr  "Male" "Male" "Male" "Male" ...
##  $ StratificationCategory2  : logi  NA NA NA NA NA NA ...
##  $ Stratification2          : logi  NA NA NA NA NA NA ...
##  $ StratificationCategory3  : logi  NA NA NA NA NA NA ...
##  $ Stratification3          : logi  NA NA NA NA NA NA ...
##  $ GeoLocation              : chr  "POINT (-92.27449074299966 34.74865012400045)" "POINT (-120.15503
##  $ ResponseID               : logi  NA NA NA NA NA NA ...
##  $ LocationID               : int  5 41 17 25 35 35 35 45 6 39 ...
##  $ TopicID                  : chr  "AST" "AST" "AST" "AST" ...
##  $ QuestionID               : chr  "AST3_1" "AST3_1" "AST4_1" "AST4_1" ...
##  $ DataValueTypeID          : chr  "NMBR" "NMBR" "NMBR" "NMBR" ...
##  $ StratificationCategoryID1: chr  "GENDER" "GENDER" "GENDER" "GENDER" ...
##  $ StratificationID1        : chr  "GENM" "GENM" "GENM" "GENM" ...
##  $ StratificationCategoryID2: logi  NA NA NA NA NA NA ...
##  $ StratificationID2        : logi  NA NA NA NA NA NA ...
##  $ StratificationCategoryID3: logi  NA NA NA NA NA NA ...
##  $ StratificationID3        : logi  NA NA NA NA NA NA ...
##  $ lat                      : num  34.7 44.6 40.5 42.3 34.5 ...
##  $ long                     : num  -92.3 -120.2 -89 -72.1 -106.2 ...
##  $ STATEFP                  : int  5 41 17 25 35 35 35 45 6 39 ...
##  $ COUNTYFP                 : int  119 69 113 27 57 57 57 79 99 89 ...
##  $ TRACTCE                  : int  4400 960100 1303 725100 963700 963700 963700 1600 1800 759000 ...
##  $ BLKGRPCE                 : int  1 2 2 2 2 2 2 1 2 1 ...
##  $ GEOID                    : num  5.12e+10 4.11e+11 1.71e+11 2.50e+11 3.51e+11 ...
##  $ OBJECTID                 : int  30558 183638 89790 112741 147838 147838 147838 196749 38466 175929
##  $ GEOID10                  : num  5.12e+10 4.11e+11 1.71e+11 2.50e+11 3.51e+11 ...
##  $ GEOID20                  : num  5.12e+10 4.11e+11 1.71e+11 2.50e+11 3.51e+11 ...
##  $ CSA                      : int  340 NA 145 148 106 106 106 192 488 198 ...
##  $ CSA_Name                 : chr  "Little Rock-North Little Rock, AR" "" "Bloomington-Pontiac, IL" "
##  $ CBSA                     : int  30780 NA 14010 49340 10740 10740 10740 17900 33700 18140 ...
##  $ CBSA_Name                : chr  "Little Rock-North Little Rock-Conway, AR" "" "Bloomington, IL" "W
##  $ CBSA_POP                 : int  734502 0 173219 938818 910012 910012 910012 816664 539301 2054062
##  $ CBSA_EMP                 : int  346204 0 86140 336137 383498 383498 383498 386734 186753 1038033
##  $ CBSA_WRK                 : int  315683 0 79904 463067 375390 375390 375390 344408 213072 970174 .
##  $ Ac_Total                 : num  427.2 784824.7 89.8 378.1 255680 ...
##  $ Ac_Water                 : num  28.8 159.2 0 0 142.4 ...
##  $ Ac_Land                  : num  398.4 784665.4 89.8 378.1 255537.6 ...
##  $ Ac_Unpr                  : num  392.5 504165 89.8 372.1 214558.9 ...
```

```
##  $ TotPop                  : int  1228 756 824 964 1464 1464 1464 769 1137 625 ...
##  $ CountHU                 : int  1260 596 296 483 991 991 991 474 618 295 ...
##  $ HH                      : int  948 355 296 373 617 617 617 356 577 228 ...
##  $ P_WrkAge                : num  0.816 0.526 0.694 0.648 0.572 0.572 0.572 0.948 0.636 0.48 ...
##  $ AutoOwn0                : int  226 0 38 9 34 34 34 16 204 123 ...
##  $ Pct_AO0                 : num  0.2384 0 0.1284 0.0241 0.0551 ...
##  $ AutoOwn1                : int  527 88 91 158 128 128 128 214 304 41 ...
##  $ Pct_AO1                 : num  0.556 0.248 0.307 0.424 0.207 ...
##  $ AutoOwn2p               : int  195 267 167 206 455 455 455 126 69 64 ...
##  $ Pct_AO2p                : num  0.206 0.752 0.564 0.552 0.737 ...
##  $ Workers                 : int  719 279 338 506 284 284 284 278 415 263 ...
##  $ R_LowWageWk             : int  154 107 108 100 84 84 84 60 101 75 ...
##  $ R_MedWageWk             : int  223 106 129 150 110 110 110 82 156 118 ...
##  $ R_HiWageWk              : int  342 66 101 256 90 90 90 136 158 70 ...
##  $ R_PCTLOWWAGE            : num  0.214 0.384 0.32 0.198 0.296 ...
##  $ TotEmp                  : int  21225 155 568 32 130 130 130 4116 9970 1 ...
##  $ E5_Ret                  : int  251 18 51 10 6 6 6 315 236 0 ...
##  $ E5_Off                  : int  11152 16 0 0 19 19 19 1444 2505 1 ...
##  $ E5_Ind                  : int  1966 58 12 1 75 75 75 245 2313 0 ...
##  $ E5_Svc                  : int  5237 59 466 20 29 29 29 964 4014 0 ...
##  $ E5_Ent                  : int  2619 4 39 1 1 1 1 1148 902 0 ...
##  $ E8_Ret                  : int  251 18 51 10 6 6 6 315 236 0 ...
##  $ E8_off                  : int  5546 1 0 0 6 6 6 868 1156 1 ...
##  $ E8_Ind                  : int  1966 58 12 1 75 75 75 245 2313 0 ...
##  $ E8_Svc                  : int  4324 7 291 4 8 8 8 951 1888 0 ...
##  $ E8_Ent                  : int  2619 4 39 1 1 1 1 1148 902 0 ...
##  $ E8_Ed                   : int  186 48 130 0 0 0 0 6 931 0 ...
##  $ E8_Hlth                 : int  727 4 45 16 21 21 21 7 1195 0 ...
##  $ E8_Pub                  : int  5606 15 0 0 13 13 13 576 1349 0 ...
##  $ E_LowWageWk             : int  3162 57 281 12 23 23 23 1193 2242 0 ...
##  $ E_MedWageWk             : int  6910 61 199 12 60 60 60 1282 3145 1 ...
##  $ E_HiWageWk              : int  11153 37 88 8 47 47 47 1641 4583 0 ...
##  $ E_PctLowWage            : num  0.149 0.368 0.495 0.375 0.177 ...
##  $ D1A                     : num  3.21014 0.00118 3.29674 1.29804 0.00462 ...
##  $ D1B                     : num  3.12862 0.0015 9.1774 2.5907 0.00682 ...
##  $ D1C                     : num  5.41e+01 3.07e-04 6.33 8.60e-02 6.06e-04 ...
##  $ D1C5_RET                : num  6.39e-01 3.57e-05 5.68e-01 2.69e-02 2.80e-05 ...
##  $ D1C5_OFF                : num  2.84e+01 3.17e-05 0.00 0.00 8.86e-05 ...
##  $ D1C5_IND                : num  5.008844 0.000115 0.133651 0.002687 0.00035 ...
##  $ D1C5_SVC                : num  1.33e+01 1.17e-04 5.19 5.37e-02 1.35e-04 ...
##  $ D1C5_ENT                : num  6.67 7.93e-06 4.34e-01 2.69e-03 4.66e-06 ...
##  $ D1C8_RET                : num  6.39e-01 3.57e-05 5.68e-01 2.69e-02 2.80e-05 ...
##  $ D1C8_OFF                : num  1.41e+01 1.98e-06 0.00 0.00 2.80e-05 ...
##   [list output truncated]
```

TODO: Insert rest of paper here

Appendix

Original data pre-processing methodology

As described in the objective section, the original data came from two sources. The disease indicators data contains location information in the form of latitude and longitude. The walkability data contains location

information in the form of Federal census location codes (FIPS codes). The pre-processing technique below was used to convert the latitude and longitude to FIPS codes, and then perform a join operation utilizing the FIPS codes. The resulting data is the original disease indicators data, augmented with the walkability information for the location corresponding to the original latitude and longitude.

In other words, for every row in the disease indicators data set, the corresponding walkability information for the region was added to that row. All of the commands are commented out to prevent them from being executed on knit since they take a long time to run.

```
#download.file("https://edg.epa.gov/EPADataCommons/public/OA/EPA_SmartLocationDatabase_V3_Jan_2021_Fina
#download.file("https://data.cdc.gov/api/views/g4ie-h725/rows.csv?accessType=DOWNLOAD", destfile="disea
```

Download the raw data

```
#walkability <- read.csv("walkability.csv")
## some of the disease data has no GeoLocation, which we cannot use for our analysis, so filter those o
#disease <- filter(read.csv("diseaseindicators.csv"), GeoLocation != "")
```

Load the data into R

```
## Extract the latitude and longitude values from the GeoLocation column using str_extract_all()
#geo_df <- str_extract_all(disease$GeoLocation, "-?[0-9]+\\.[0-9]+")

## Convert the extracted values to numeric and assign them to the corresponding latitude and longitude
#disease$lat <- as.numeric(sapply(geo_df, function(x) x[2]))
#disease$long <- as.numeric(sapply(geo_df, function(x) x[1]))
```

Extract the latitude and longitude into separate columns

Fetch the geographic information required to map latitude and logitude to FIPS blocks The tigris library provides a function "block_groups" which returns geographic information about every FIPS block. This geographic information can be used to convert latitude and longitude to FIPS block. The following code downloads all of the block_groups for every block in the walkability data set.

```
## create data frame for block_groups data
#allblockgroups <- data.frame(matrix(ncol=6, nrow=0))
#colnames(allblockgroups) <- c('STATEFP', 'COUNTYFP', 'TRACTCE', 'BLKGRPCE', 'GEOID', 'geometry')

## get block geography data for each state in the walkability dataset
#stateCodes <- data.frame(unique(walkability$STATEFP))
#for (i in 1:nrow(stateCodes)) {
#  stateCode=stateCodes[[1]][i]
#  counties = distinct(filter(walkability, STATEFP == stateCode), COUNTYFP)$COUNTYFP
#  new_blocks <- block_groups(state=stateCodes[[1]][i], counties) %>%
```

```
#    select(STATEFP, COUNTYFP, TRACTCE, BLKGRPCE, GEOID, geometry)
#  allblockgroups <- rbind(allblockgroups, new_blocks)
#}
```

```
#my_points <- data.frame(
#  x = disease$lat,
#  y = disease$long
#) %>%
#  st_as_sf(coords = c("y", "x"),
#     crs = st_crs(allblockgroups))

#my_points_blocks <- st_join(my_points, allblockgroups)
#disease$STATEFP = as.integer(my_points_blocks$STATEFP)
#disease$COUNTYFP = as.integer(my_points_blocks$COUNTYFP)
#disease$TRACTCE = as.integer(my_points_blocks$TRACTCE)
#disease$BLKGRPCE = as.integer(my_points_blocks$BLKGRPCE)
#disease$GEOID = as.numeric(my_points_blocks$GEOID)
```

Use block geographies to convert longitude and latitude to FIPS blocks

```
# Join the disease data with the walkability data
#disease_with_walkability <- left_join(disease, walkability,
#                                      by = c("STATEFP", "COUNTYFP", "TRACTCE", "BLKGRPCE"))
```

Join the disease indicators and walkability data sets based on FIPS blocks

```
#write.csv(disease_with_walkability, file = "disease_with_walkability.csv")
```

Export the joined data to be used for further processing later.