ERHS535\_Homework5

Mollie Uhrig

2024-11-18

List of files in data directory:

getwd()

## [1] "/Users/mollieuhrig/Documents/Classes/Fall 2024 Classes/ERHS 535/Homework/Homework 5/ERHS535\_Homework5/writing"

list.files("../data")

## [1] "homicide-data.csv"

Reading in data:

#Reading data from Github  
library(readr)  
  
hom\_url <- paste0("https://raw.githubusercontent.com/washingtonpost/data-homicides/refs/heads/master/homicide-data.csv")  
  
homicides <- read\_csv(hom\_url)

## Rows: 52179 Columns: 12  
## ── Column specification ────────────────────────────────────────────────────────  
## Delimiter: ","  
## chr (9): uid, victim\_last, victim\_first, victim\_race, victim\_age, victim\_sex...  
## dbl (3): reported\_date, lat, lon  
##   
## ℹ Use `spec()` to retrieve the full column specification for this data.  
## ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

Cleaning data:

#Organizing the data to narrow down the dataset for relevance.   
library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.4 ✔ purrr 1.0.2  
## ✔ forcats 1.0.0 ✔ stringr 1.5.1  
## ✔ ggplot2 3.5.1 ✔ tibble 3.2.1  
## ✔ lubridate 1.9.3 ✔ tidyr 1.3.1  
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(dplyr)  
homicides\_co <- homicides %>%   
 mutate(city\_name = paste0(city, ", ", state)) %>%   
 filter(city\_name == "Denver, CO") %>%   
 select("lat", "lon", "victim\_race", "disposition")

Separating categories of homicides:

#Categories of homicides- unsolved and solved  
homicides\_co\_sep <- homicides\_co %>%   
 mutate(status = ifelse (disposition == "Closed without arrest" | disposition == "Open/No arrest", "Unsolved", "Solved"))

Determining three races with highest number of homicides:

#Using fct\_lump to determine races with highest number of homicides and make an "other" group.   
library(forcats)  
library(dplyr)  
homicides\_co\_sep <- homicides\_co\_sep%>%   
 mutate(victim\_race = forcats::fct\_lump(victim\_race, n = 3))

Denver neighborhoods:

#Reading in data for Denver block groups  
library(tigris)

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

denver <- block\_groups(state = c("Colorado"), county = c("Denver"), cb= TRUE, class= "sf")

## Retrieving data for the year 2022

## | | | 0% | |= | 1% | |== | 3% | |=== | 4% | |=== | 5% | |===== | 7% | |===== | 8% | |====== | 8% | |======= | 9% | |========== | 14% | |=========== | 16% | |============ | 18% | |============= | 19% | |============== | 20% | |=============== | 22% | |================= | 24% | |================== | 26% | |==================== | 29% | |===================== | 30% | |================================================ | 68% | |================================================= | 70% | |=================================================== | 72% | |==================================================== | 74% | |===================================================== | 76% | |====================================================== | 77% | |======================================================= | 79% | |======================================================== | 81% | |=========================================================== | 84% | |============================================================ | 86% | |============================================================== | 89% | |================================================================ | 91% | |================================================================== | 94% | |=================================================================== | 96% | |======================================================================| 100%

Setting up homicides data in sf format:

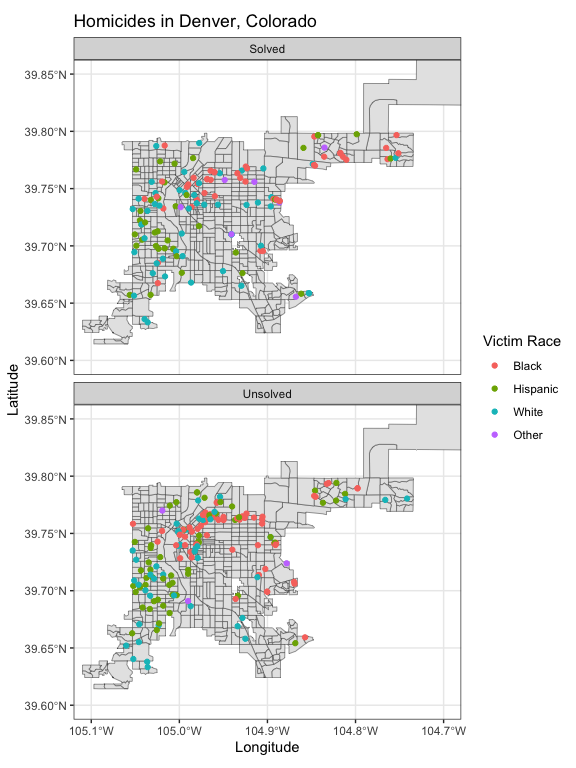
library(sf)

## Linking to GEOS 3.11.0, GDAL 3.5.3, PROJ 9.1.0; sf\_use\_s2() is TRUE

#Setting up homicides in sf format  
hom\_co\_sf <- homicides\_co\_sep %>%   
 st\_as\_sf(coords = c("lon", "lat"), crs=4326)

Making the map (option 1):

#Layering the data for Denver blocks with the homicides coordinates and generating the map.   
  
library(sf)  
library(ggplot2)  
  
ggplot() +  
 geom\_sf(data = denver) +  
 geom\_sf(data = hom\_co\_sf, aes(color= victim\_race)) +  
 coord\_sf(datum = st\_crs(4326)) +  
 labs(y="Latitude", x="Longitude", color= "Victim Race")+  
 ylim(39.6, 39.85)+  
 xlim(105.1, 104.7)+  
 ggtitle("Homicides in Denver, Colorado") +  
 theme\_bw()+  
 facet\_wrap(~ status, ncol = 1)



#Used crs of 4326 for standard coordinate data.   
#Facet\_wrap with ncol=1 to make the maps larger and easier to see.