Homework\_5

Taru Dutt

11/15/2021

## R Markdown

## Loading packages

library(rmarkdown)  
library(tidyverse)  
library(forcats)  
library(lubridate)  
library(tigris)  
library(sf)  
options(tigris\_class = "sf")  
options(tigris\_use\_cache = TRUE)

## Loading data

homicide\_data <- read\_csv("../data/homicide-data.csv")  
homicide\_data

## # A tibble: 52,179 × 12  
## uid reported\_date victim\_last victim\_first victim\_race victim\_age  
## <chr> <dbl> <chr> <chr> <chr> <chr>   
## 1 Alb-000001 20100504 GARCIA JUAN Hispanic 78   
## 2 Alb-000002 20100216 MONTOYA CAMERON Hispanic 17   
## 3 Alb-000003 20100601 SATTERFIELD VIVIANA White 15   
## 4 Alb-000004 20100101 MENDIOLA CARLOS Hispanic 32   
## 5 Alb-000005 20100102 MULA VIVIAN White 72   
## 6 Alb-000006 20100126 BOOK GERALDINE White 91   
## 7 Alb-000007 20100127 MALDONADO DAVID Hispanic 52   
## 8 Alb-000008 20100127 MALDONADO CONNIE Hispanic 52   
## 9 Alb-000009 20100130 MARTIN-LEYVA GUSTAVO White 56   
## 10 Alb-000010 20100210 HERRERA ISRAEL Hispanic 43   
## # … with 52,169 more rows, and 6 more variables: victim\_sex <chr>, city <chr>,  
## # state <chr>, lat <dbl>, lon <dbl>, disposition <chr>

## Cleaning data

For cleaning data: 1. I am uniting city and state column as “city\_state” 2. Setting date class for reported date 3. Changing names of victim first and victim last columns to sentence as all letters are in uppercase. 4.

homicide\_data\_atlanta <- homicide\_data %>%   
 unite(col = "city\_state", city, state, sep = ",") %>%   
 mutate(victim\_first = str\_to\_sentence(victim\_first),  
 victim\_last = str\_to\_sentence(victim\_last),  
 disposition = str\_to\_lower(disposition),  
 disposition = str\_replace(disposition, "closed without arrest", "unsolved"),  
 disposition = str\_replace(disposition, "open/no arrest", "unsolved"),  
 disposition = str\_replace(disposition, "closed by arrest", "solved")) %>%   
 filter(city\_state == "Atlanta,GA") %>%   
 print()

## # A tibble: 973 × 11  
## uid reported\_date victim\_last victim\_first victim\_race victim\_age  
## <chr> <dbl> <chr> <chr> <chr> <chr>   
## 1 Atl-000756 20070110 Birdsong Ernest Black 29   
## 2 Atl-000757 20070110 Reed Michael Black 29   
## 3 Atl-000758 20070114 Feliciano Johnny Hispanic 36   
## 4 Atl-000759 20070114 Spencer Charles Black 26   
## 5 Atl-000760 20070120 Pope Luke Black 75   
## 6 Atl-000761 20070128 Leary Ernest Black 22   
## 7 Atl-000762 20070204 Hunt Bruce Black 40   
## 8 Atl-000763 20070215 Turner Deandra Black 1   
## 9 Atl-000764 20070225 Bateman Demario Black 25   
## 10 Atl-000765 20070225 Mills Terry Black 25   
## # … with 963 more rows, and 5 more variables: victim\_sex <chr>,  
## # city\_state <chr>, lat <dbl>, lon <dbl>, disposition <chr>

## Creating map

# Adding race parameter  
  
homicide\_data\_atlanta$victim\_race <- as.factor(homicide\_data\_atlanta$victim\_race)   
  
  
# Creating sf   
  
  
sf\_homicide\_data\_atlanta <- st\_as\_sf(x = homicide\_data\_atlanta,   
 coords = c("lon", "lat")) %>%   
 st\_set\_crs(4269)   
  
# Plot graph  
  
atlanta <- tracts("GA", county = "fulton", cb = TRUE)  
  
ggplot() +  
 geom\_sf(data = atlanta) +  
 geom\_sf(data = sf\_homicide\_data\_atlanta, alpha = 0.5,   
 aes(color = fct\_lump(victim\_race, n = 3))) +  
 facet\_wrap( ~ disposition) +  
 scale\_color\_viridis\_d() +  
 ggtitle("Atlanta homicide incidents",   
 subtitle = "Separated by solved or unsolved homicides") +  
 scale\_color\_discrete(name = 'Divided by Race') +  
 theme(axis.text.x = element\_blank(),   
 axis.text.y = element\_blank(),  
 plot.title = element\_text(color="black", size=20, face="bold"),  
 strip.text = element\_text(size=14, face = "bold"))

