Homework5

Aeriel, Rebecca, and Grant

11/16/2018

## Libraries

library(readr)  
library(dplyr)  
library(sf)  
library(tigris)  
library(acs)  
library(tidyr)  
library(scales)  
library(ggplot2)  
library(knitr)  
library(ggthemes)  
library(forcats)  
library(lubridate)  
library(tidyr)  
library(knitr)

## Load Data

homicides <- read\_csv('data/homicide-data.csv')  
  
colnames(homicides)

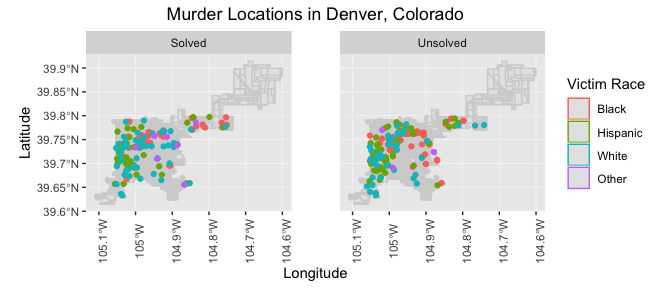
## Plot 1

### Rebecca

Pick one city in the data (DENVER!!!). Create a map showing the locations of the homicides in that city, using the sf framework discussed in class. Use tigris to download boundaries for some sub-city geography (e.g., tracts, block groups, county subdivisions) to show as a layer underneath the points showing homicides. Use different facets for solved versus unsolved homicides and different colors to show the three race groups with the highest number of homicides for that city (you may find the fct\_lump function from forcats useful for this).

## Question 1  
##Select for Denvver and convert to just solved or unsolved  
denver <- homicides %>%   
 subset(city == "Denver") %>%   
 mutate(disposition = recode(disposition,   
 "Closed without arrest" = 'Unsolved',  
 "Open/No arrest"= 'Unsolved',  
 "Closed by arrest" = 'Solved'))  
  
##Create sub-city layer by census blocks  
api.key.install(key="a22c251a387f5312b9e5c2e5ad21494088e984f6") ##not sure if I need an api for this  
denver\_map <- blocks(state = "CO", county = "Denver County", class= "sf")

##Change coordinates to sf object and lump together the top three victim\_race  
denver\_homicides <- st\_as\_sf(denver, coords= c("lon", "lat")) %>%   
 st\_set\_crs(4269) %>%   
 mutate(victim\_race = forcats::fct\_lump(victim\_race, n=3))  
  
##Plot map with two facets for Solved and Unsolved disposition, colors for race, and pretty axis. Not sure if I can get rid of the background noise   
ggplot() +   
 geom\_sf(data = denver\_map, color = "lightgray") +   
 geom\_sf(data = denver\_homicides, aes(color= victim\_race))+  
 facet\_grid(~disposition) +  
 ggtitle("Murder Locations in Denver, Colorado")+  
 theme(plot.title = element\_text(hjust=0.5))+  
 labs( y= "Latitude", x= "Longitude", color= "Victim Race") +  
 theme(panel.spacing = unit(.5, unit="in")) +  
 theme(axis.text.x = element\_text(angle=90))



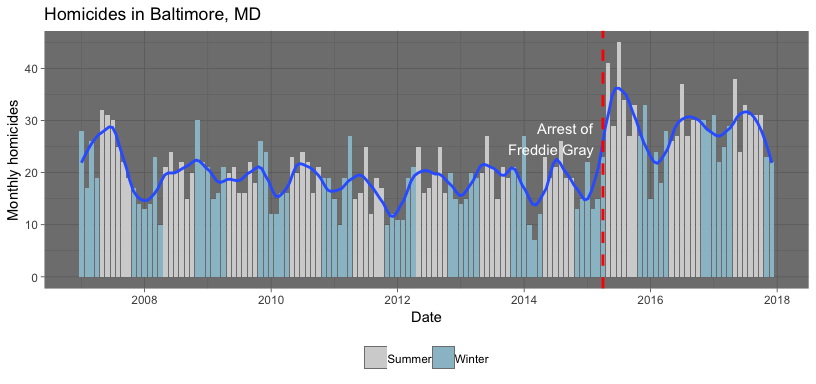
## Plot 2

### Aeriel

Recreate the graph shown below. It shows monthly homicides in Baltimore, with a reference added for the date of the arrest of Freddie Gray and color used to show colder months (November through April) versus warmer months (May through October). There is a smooth line added to help show seasonal and long-term trends in this data.

freddie <- homicides %>%   
 filter(victim\_last == "GREY" & victim\_first == "FREDDIE CARLOS") %>%   
 mutate(reported\_date = ymd(reported\_date),  
 month = month(reported\_date, label = TRUE),  
 year = year(reported\_date),  
 Date = format(reported\_date, "%Y-%m", label = TRUE),  
 Date = ymd(Date, truncated = 1)) %>%   
 group\_by(Date) %>%   
 count()  
  
homicides %>%   
 filter(city == "Baltimore") %>%   
 mutate(reported\_date = ymd(reported\_date),  
 month = month(reported\_date, label = TRUE),  
 year = year(reported\_date),  
 Date = format(reported\_date, "%Y-%m", label = TRUE),  
 Date = ymd(Date, truncated = 1),  
 Summer = month %in% "May" | month %in% "Jun" | month %in% "Jul" |  
 month %in% "Aug" | month %in% "Sep" | month %in% "Oct",  
 Winter = Summer == FALSE,  
 Summer = ifelse(Summer == TRUE, "Summer", ""),  
 Winter = ifelse(Winter == TRUE, "Winter", "")) %>%   
 unite(season, Summer, Winter, sep = "") %>%   
 group\_by(Date, season) %>%   
 count() %>%   
 ungroup() %>%   
 ggplot(aes(x = Date, y = n)) +  
 geom\_bar(aes(fill = season), stat = "identity") +  
 geom\_smooth(se = FALSE, n = 200, span = 0.1) +  
 geom\_vline(data = freddie, aes(xintercept = Date),   
 color = "red", linetype = "dashed", size = 1) +  
 geom\_text(data = freddie, label = "Arrest of \n Freddie Gray",   
 color = "white", vjust = -3.7, hjust = 1.1) +  
 scale\_fill\_manual(values = c("Winter" = "lightblue3",   
 "Summer" = "lightgray")) +  
 labs(x = "Date", y = "Monthly homicides") +  
 ggtitle("Homicides in Baltimore, MD") +  
 theme\_dark() +  
 theme(legend.title = element\_blank(),  
 legend.position = "bottom")

## `geom\_smooth()` using method = 'loess' and formula 'y ~ x'



## Plot 3

### Grant

1. Create one more plot using this data. Work with your group to create a plot that follows the principles of good plotting and that you think illustrates something interesting in the data. Write a paragraph explaining what the plot is showing and why you find it interesting.

homicides2 <- homicides %>%  
 filter(city == "Chicago" | city == "Denver") %>%   
 mutate(reported\_date = ymd(reported\_date)) %>%   
 separate(reported\_date, c("year", "month", "day", sep = "-")) %>%   
 mutate(day = 15) %>%   
 unite(date, c(day, month, year), sep = "-", remove = FALSE) %>%   
 mutate(date = as.Date(dmy(date))) %>%   
 select(victim\_age, victim\_race, victim\_sex, city, date) %>%   
 group\_by(city, date) %>%   
 count()

## Warning: Expected 4 pieces. Missing pieces filled with `NA` in 5847  
## rows [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19,  
## 20, ...].

class(homicides2$date)

## [1] "Date"

chic\_vs\_denplot <- homicides2 %>%   
 ggplot(aes(x = date, y = n))+  
 geom\_col()+  
 facet\_grid(city ~ .)+  
 theme\_few()+  
 labs(x = "Date", y = "Monthly Homicides",   
 title = "Chicago vs. Denver Homicides")+  
 geom\_smooth(size = 1, weight = 0.5, color = "Red")  
chic\_vs\_denplot

## `geom\_smooth()` using method = 'loess' and formula 'y ~ x'

