HW\_5\_B

Jered May

2024-11-17

#Load Data and Libraries

library(dplyr)  
library(tidyr)  
library(stringr)  
library(readr)  
library(purrr)  
library(broom)  
library(ggplot2)  
library(scales)  
library(forcats)  
library(lubridate)  
  
homicides\_url <- paste0("https://raw.githubusercontent.com/washingtonpost/",  
 "data-homicides/refs/heads/master/homicide-data.csv")  
  
homicides <- read\_csv(homicides\_url) %>%  
 unite(city\_name, city, state, sep = ",", remove =FALSE)  
  
  
homicides$reported\_date <- as.Date(as.character(homicides$reported\_date),  
 format = "%Y%m%d")  
  
# confirm reported\_date is a Date class  
print(class(homicides$reported\_date))

## [1] "Date"

# Debuging

# Debugging   
tulsa\_records <- homicides %>%  
 filter(city == "Tulsa",  
 victim\_last == "HARRIS")  
print(tulsa\_records, width = Inf)

## # A tibble: 7 × 13  
## uid reported\_date victim\_last victim\_first victim\_race victim\_age  
## <chr> <date> <chr> <chr> <chr> <chr>   
## 1 Tul-000236 2007-11-14 HARRIS JARRED White 21   
## 2 Tul-000321 2009-04-28 HARRIS COLBY Black 20   
## 3 Tul-000565 2013-05-25 HARRIS RONALD Black 23   
## 4 Tul-000612 2014-04-09 HARRIS STEVEN Black 39   
## 5 Tul-000673 2015-03-12 HARRIS DEMARIO Black 34   
## 6 Tul-000691 2015-07-12 HARRIS JULIAN Black 33   
## 7 Tul-000769 2016-09-30 HARRIS LESLIE White 41   
## victim\_sex city\_name city state lat lon disposition   
## <chr> <chr> <chr> <chr> <dbl> <dbl> <chr>   
## 1 Male Tulsa,OK Tulsa OK 36.1 -95.9 Closed by arrest  
## 2 Male Tulsa,OK Tulsa OK 36.1 -96.0 Closed by arrest  
## 3 Male Tulsa,OK Tulsa OK 36.1 -96.0 Closed by arrest  
## 4 Male Tulsa,OK Tulsa OK 36.2 -96.0 Closed by arrest  
## 5 Male Tulsa,OK Tulsa OK 36.1 -95.9 Closed by arrest  
## 6 Male Tulsa,OK Tulsa OK 36.1 -95.9 Open/No arrest   
## 7 Female Tulsa,AL Tulsa AL 36.1 -96.0 Closed by arrest

# mutating the one record of Tulsa, AL  
homicides <- read\_csv(homicides\_url) %>%  
 mutate(state = if\_else(city == "Tulsa", "OK", state)) %>%  
 unite(city\_name, city, state, sep =", ", remove = FALSE)  
  
# Check to see that the single Tulsa, AL record is changed  
tulsa\_records <- homicides %>%  
 filter(city == "Tulsa",  
 victim\_last == "HARRIS")  
print(tulsa\_records, width = Inf)

## # A tibble: 7 × 13  
## uid reported\_date victim\_last victim\_first victim\_race victim\_age  
## <chr> <dbl> <chr> <chr> <chr> <chr>   
## 1 Tul-000236 20071114 HARRIS JARRED White 21   
## 2 Tul-000321 20090428 HARRIS COLBY Black 20   
## 3 Tul-000565 20130525 HARRIS RONALD Black 23   
## 4 Tul-000612 20140409 HARRIS STEVEN Black 39   
## 5 Tul-000673 20150312 HARRIS DEMARIO Black 34   
## 6 Tul-000691 20150712 HARRIS JULIAN Black 33   
## 7 Tul-000769 20160930 HARRIS LESLIE White 41   
## victim\_sex city\_name city state lat lon disposition   
## <chr> <chr> <chr> <chr> <dbl> <dbl> <chr>   
## 1 Male Tulsa, OK Tulsa OK 36.1 -95.9 Closed by arrest  
## 2 Male Tulsa, OK Tulsa OK 36.1 -96.0 Closed by arrest  
## 3 Male Tulsa, OK Tulsa OK 36.1 -96.0 Closed by arrest  
## 4 Male Tulsa, OK Tulsa OK 36.2 -96.0 Closed by arrest  
## 5 Male Tulsa, OK Tulsa OK 36.1 -95.9 Closed by arrest  
## 6 Male Tulsa, OK Tulsa OK 36.1 -95.9 Open/No arrest   
## 7 Female Tulsa, OK Tulsa OK 36.1 -96.0 Closed by arrest

# Filter

# filtering for Baltimore homicides with seasons  
Baltimore\_homicides <- homicides %>%  
 filter(city\_name == "Baltimore, MD") %>%  
 mutate(  
 reported\_date = ymd(reported\_date),  
 year = year(reported\_date),  
 month = month(reported\_date),  
 season = case\_when(  
 month %in% c(05, 06, 07, 08, 09, 10) ~ "Summer",  
 TRUE ~ "Winter"  
 )  
 )  
  
#Counts  
monthly\_counts <- Baltimore\_homicides %>%  
 group\_by(year, month, season) %>%  
 nest() %>%  
 mutate(  
 count = map\_int(data, nrow)  
 ) %>%  
 ungroup() %>%  
 mutate(  
 date = make\_date(year, month, 1)  
 ) %>%  
 select(year, month, count, season, date)  
  
  
print(monthly\_counts)

## # A tibble: 132 × 5  
## year month count season date   
## <dbl> <dbl> <int> <chr> <date>   
## 1 2007 1 28 Winter 2007-01-01  
## 2 2007 2 17 Winter 2007-02-01  
## 3 2007 3 26 Winter 2007-03-01  
## 4 2007 4 19 Winter 2007-04-01  
## 5 2007 5 32 Summer 2007-05-01  
## 6 2007 6 31 Summer 2007-06-01  
## 7 2007 7 30 Summer 2007-07-01  
## 8 2007 8 25 Summer 2007-08-01  
## 9 2007 9 22 Summer 2007-09-01  
## 10 2007 10 19 Summer 2007-10-01  
## # ℹ 122 more rows

#Graph

# Plot with ggplot2  
ggplot(monthly\_counts, aes(x = date, y = count, fill = season)) +  
 geom\_bar(stat = "identity", position = "dodge") +  
 scale\_fill\_manual(values = c("Summer" = "gray", "Winter" = "lightblue")) +  
 geom\_vline(xintercept = as.Date("2015-04-01"), linetype = "longdash",  
 color = "red", size = 1) +  
 annotate("text",   
 x = as.Date("2015-04-12"),   
 y = max(monthly\_counts$count) - 19,   
 label = ("Arrest of\nFreddie Gray"),   
 color = "white",   
 hjust = 1.13,  
 vjust = -1) +  
 labs(  
 x = "Date",  
 y = "Monthly homicides",  
 title = "Homicides in Baltimore, MD") +  
 theme\_dark() +  
 scale\_y\_continuous(  
 limits = c(0, max(monthly\_counts$count))) +  
 theme(legend.position = "bottom") +  
 guides(fill = guide\_legend(title = NULL)) +  
 geom\_smooth(aes(group = 1), method = "loess", color = "blue", span = 0.1,  
 se = FALSE, show.legend =FALSE)

