Homework_5

Andrea_Sanchez_Hidalgo

2025-12-01

```r
# Create README.md file from R

readme_text <- "

# Homework 5

This README file contain instructions for HW5.

## Description
The purpose of this HW is learn to commmit to GIthub in R.

## Requirements or Packages

#library(dplyr)
#library(ggplot2)
#library(tidyr)
#library(tidyverse)
#library(lubridate)

## Bibliography
Some information was taken from Wikipedia and credit was given.

"

# Write the file to the project root
writeLines(readme_text, 'README.md')
```

```r
##It shows monthly homicides in Baltimore, with a reference added for the date of the arrest of Freddie

##Library load
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(ggplot2)
library(tidyr)
library(tidyverse)
```

```
## Warning: package 'tibble' was built under R version 4.4.3

## Warning: package 'readr' was built under R version 4.4.3

## Warning: package 'stringr' was built under R version 4.4.3

## Warning: package 'forcats' was built under R version 4.4.3

## Warning: package 'lubridate' was built under R version 4.4.3

## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v forcats   1.0.1      v readr     2.1.5
## v lubridate 1.9.4      v stringr   1.5.2
## v purrr     1.0.4      v tibble    3.3.0

## -- Conflicts ------------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```r
library(lubridate)

##read file
getwd()
```

```
## [1] "C:/Users/Andrea/OneDrive - Colostate/Documentos/PhD Classes/Fall 2025/EHRS535/EHRS-ASH"
```

```r
homicides <- read.csv ("https://raw.githubusercontent.com/AndreaSH15/EHRS-ASH/refs/heads/main/homicide-d

View(homicides)
str(homicides)
```

```
## 'data.frame':     52179 obs. of  12 variables:
##  $ uid          : chr  "Alb-000001" "Alb-000002" "Alb-000003" "Alb-000004" ...
##  $ reported_date: int  20100504 20100216 20100601 20100101 20100102 20100126 20100127 20100127 20100:
##  $ victim_last  : chr  "GARCIA" "MONTOYA" "SATTERFIELD" "MENDIOLA" ...
##  $ victim_first : chr  "JUAN" "CAMERON" "VIVIANA" "CARLOS" ...
##  $ victim_race  : chr  "Hispanic" "Hispanic" "White" "Hispanic" ...
##  $ victim_age   : chr  "78" "17" "15" "32" ...
##  $ victim_sex   : chr  "Male" "Male" "Female" "Male" ...
##  $ city         : chr  "Albuquerque" "Albuquerque" "Albuquerque" "Albuquerque" ...
##  $ state        : chr  "NM" "NM" "NM" "NM" ...
##  $ lat          : num  35.1 35.1 35.1 35.1 35.1 ...
##  $ lon          : num  -107 -107 -107 -107 -107 ...
##  $ disposition  : chr  "Closed without arrest" "Closed by arrest" "Closed without arrest" "Closed by
```

```r
##extract infromation from Baltimore
homicides_baltimore <- homicides %>%
  filter(city == "Baltimore") %>%
  mutate(
    reported_date = ymd(reported_date),
    year = year(reported_date),
    month = month(reported_date),
    season = if_else(month %in% c(11,12,1,2,3,4), "colder", "warmer")
  ) %>%
  group_by(year, month, season, disposition) %>%
  summarise(homicides = n(), .groups = "drop")

homicides_baltimore_monthly <- homicides_baltimore %>%
  group_by(year, month, season) %>%
  summarise(total_homicides = sum(homicides), .groups = "drop")

View(homicides_baltimore)
```

Freddie Gray

Interestingly, this database does not contain the name of Freddy Gray as a victim, however, according to Wikipedia, he was arrested on April 12, 2015, for possession of a knife, however, he was attacked by police officers on the same date and died in R Adams Cowley Shock Trauma Center. Gray's death on April 19, 2015, was ascribed to injuries to his cervical spinal cord.

Source: https://en.wikipedia.org/wiki/Killing_of_Freddie_Gray

Note: As he was not in the data set obtained, the date of the arrest used on this exercise (April 12, 2015) was taken from Wikipedia.

```r
##Fix year of dataset
homicides_plotdata <- homicides_baltimore %>%
  mutate(
    date = ymd(paste(year, month, "01", sep = "-"))
  ) %>%
  group_by(date, season) %>%
  summarise(homicides = sum(homicides), .groups = "drop")

##Check
View(homicides_plotdata)

# Wikipedia Freddy Gray arrest
freddy_date <- ymd("2015-04-12")

##Graph

ggplot(homicides_plotdata, aes(x = date, y = homicides)) +
  geom_col(aes(fill = season), alpha = 0.75) +
  geom_smooth(se = FALSE, color = "blue", size = 1.0, span = 0.1) +
  scale_fill_manual(values = c("colder" = "lightblue", "warmer" = "grey70")) +
  scale_x_date(date_breaks = "1 year", date_labels = "%Y") +
  labs(
    title = "Homicides in Baltimore",
    x = "Year",
    y = "Monthly homicides",
```

```
    fill = ""
  ) +
  theme_minimal(base_size = 13) +  # first apply theme_minimal
  theme(
    panel.background = element_rect(fill = "grey40", color = NA), # now this works
    legend.position = "bottom",
    panel.grid.minor = element_blank(),
    panel.grid.major.x = element_blank()
  ) +
  geom_vline(xintercept = freddy_date, color = "red", linetype = "dashed", size = 1) +
  # Add text label
  annotate("text", x = freddy_date, y = max(homicides_plotdata$homicides) * 0.95,
           label = "Arrest of Freddy Gray", color = "white", angle = 0, vjust = -0.1)
```
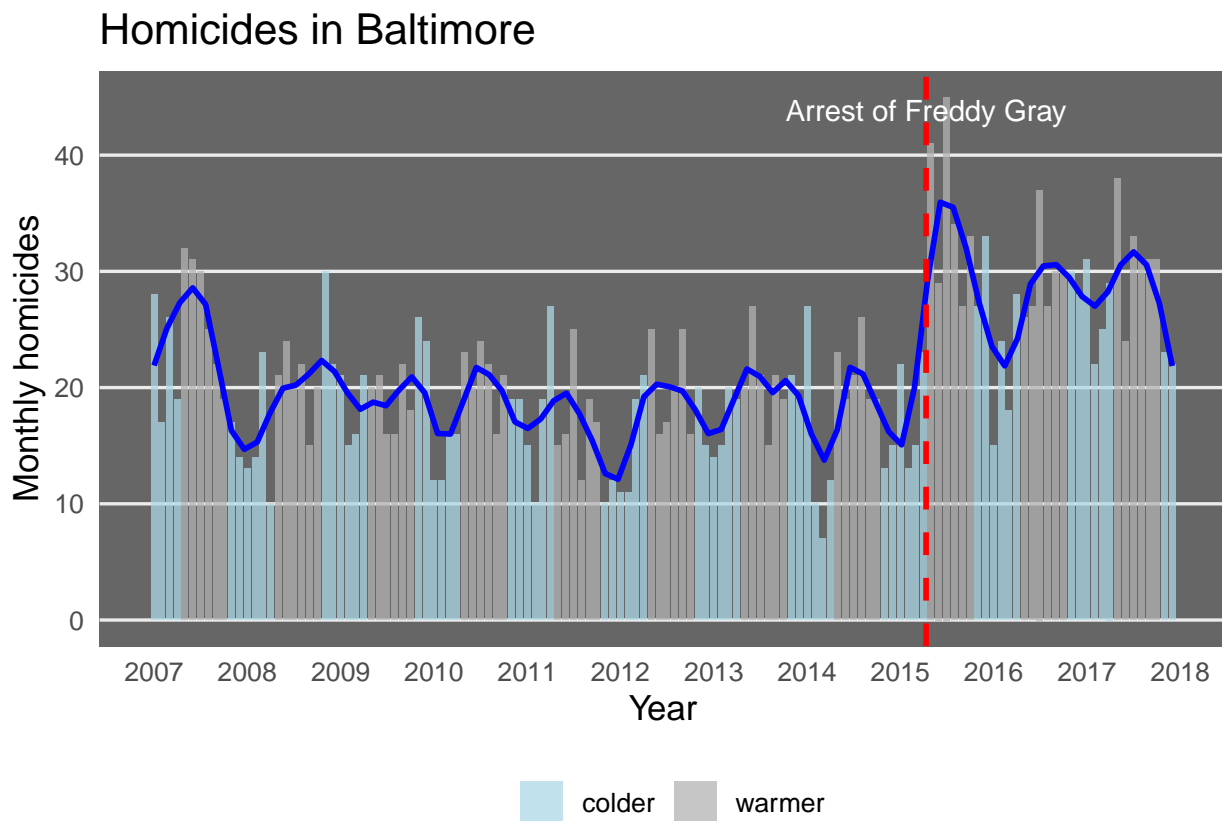
```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

```
## 'geom_smooth()' using method = 'loess' and formula = 'y ~ x'
```



```
```