HW_5

Kalex Robledo

2025-11-17

Read in homicide dataset

```
#Set working directory to project root
setwd("~/R_Programming/HW_5")
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.4.3
```

```
library(dplyr)
```

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

```
##
## 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
```

```
library(lubridate)
```

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

```
##
## Attaching package: 'lubridate'
```

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

```r
# Set global chunk options
knitr::opts_chunk$set(echo = TRUE, warning = FALSE, message = FALSE)
wp_data <- read.csv("data/homicide-data.csv")
head(wp_data)
```

```
##          uid reported_date victim_last victim_first victim_race victim_age
## 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
##   victim_sex        city state      lat       lon           disposition
## 1       Male Albuquerque    NM 35.09579 -106.5386 Closed without arrest
## 2       Male Albuquerque    NM 35.05681 -106.7153      Closed by arrest
## 3     Female Albuquerque    NM 35.08609 -106.6956 Closed without arrest
## 4       Male Albuquerque    NM 35.07849 -106.5561      Closed by arrest
## 5     Female Albuquerque    NM 35.13036 -106.5810 Closed without arrest
## 6     Female Albuquerque    NM 35.15111 -106.5378        Open/No arrest
```

Clean + Prep to only include Baltimore observations

```r
baltimore <- wp_data %>%
  filter(city == "Baltimore") %>%
  mutate(
    reported_date = as.Date(as.character(reported_date), format = "%Y%m%d"),
    year = year(reported_date),
    month = month(reported_date),
    season = ifelse(month %in% c(5,6,7,8,9,10), "Summer", "Winter")
  )
```

Aggregate monthly counts

```r
baltimore_monthly <- baltimore %>%
  mutate(month_date = floor_date(reported_date, "month")) %>%
  group_by(month_date, season) %>%
  summarise(n = n(), .groups = "drop")

# Freddie Gray arrest date
freddie_gray_date <- as.Date("2015-04-12")
```

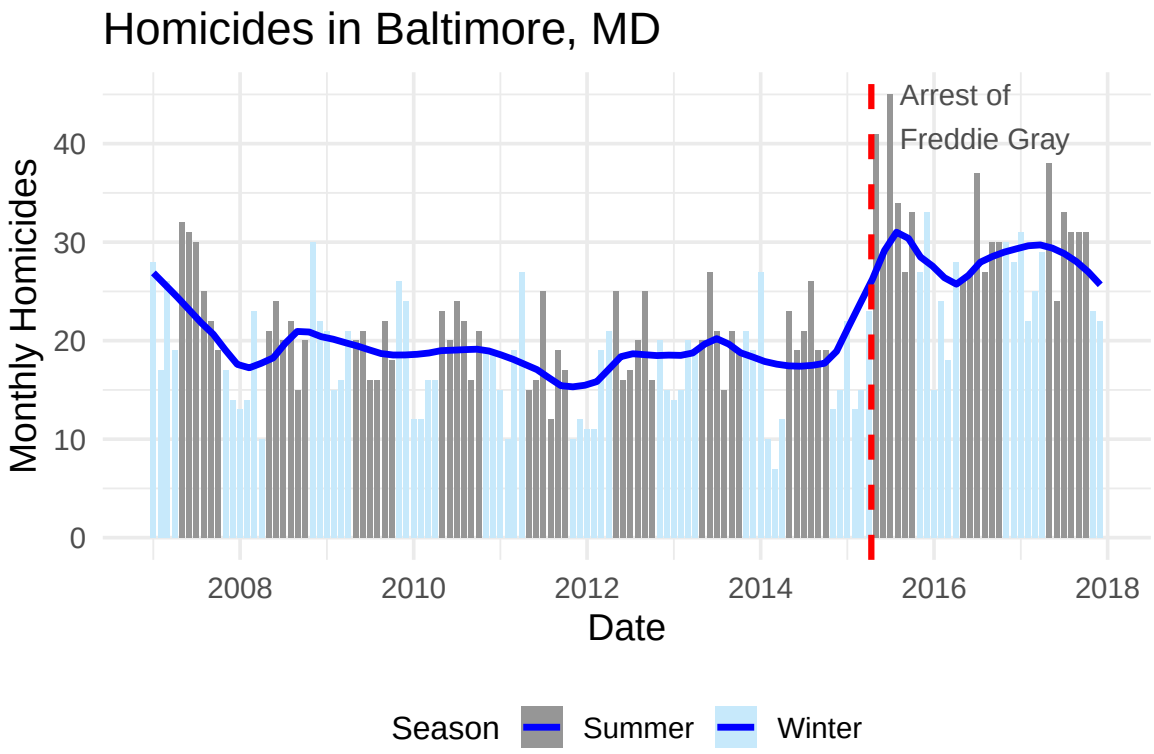
Bar plot:

```
baltimore_plot <- ggplot(baltimore_monthly, aes(x = month_date, y = n, fill = season)) +
  geom_col(width = 25) +  # bars with fill for legend
  scale_fill_manual(values = c("Summer" = "#969696", "Winter" = "#c7e9fb")) +
  geom_smooth(aes(group = 1), se = FALSE, color = "blue", size = 1.2,
              method = "loess", span = 0.2) +  # tighter trend line
  geom_vline(xintercept = as.numeric(freddie_gray_date),
             color = "red", linetype = "dashed", size = 1.1) +
  annotate("text",
           x = freddie_gray_date + 120,
           y = max(baltimore_monthly$n) * 0.95,
           label = "Arrest of\nFreddie Gray",
           color = "gray30",
           hjust = 0,
           size = 4) +
  labs(
 title = "Homicides in Baltimore, MD",
    x = "Date",
    y = "Monthly Homicides",
    fill = "Season"  # legend title
  ) +
  theme_minimal(base_size = 14) +
  theme(
    legend.position = "bottom",          # move legend below x-axis
    legend.title = element_text(size = 12),
    legend.text = element_text(size = 11),
    plot.margin = margin(20, 20, 20, 20)
  )
```

Save the figure

```
ggsave("figures/baltimore_monthly_homicides.png",
       baltimore_plot, width = 10, height = 4.5)

baltimore_plot
```


Homicides in Baltimore, MD
Arrest of
Freddie Gray
Monthly Homicides
0
10
20
30
40
2008
2010
2012
2014
2016
2018
Date
Season
Summer
Winter