

HWK #5 – Homicides

Melanie Luu

1. Introduction

For Homework #5, I will be working with the Washington Post homicide dataset. This document will eventually contain all analysis, plots, and interpretation, but for the first commit it only includes the structure and setup.

2. Project Setup Summary

- Created an R Project for HWK #5
 - Added required folders:
 - `data/`
 - `figures/`
 - `writing/`
 - Created this RMarkdown document inside the `writing/` directory
 - Initialized git for the project
 - Ready to make the first commit
-

3. Load Libraries (placeholder)

4. Placeholder: Data Import

```
# Read data from the `data` subdirectory (relative to project root)
homicides_raw <- readr::read_csv("data/homicide-data.csv")

# Glimpse at the structure
dplyr::glimpse(homicides_raw)

## Rows: 52,179
## Columns: 12
## $ uid      <chr> "Alb-000001", "Alb-000002", "Alb-000003", "Alb-000004", ~
## $ reported_date <dbl> 20100504, 20100216, 20100601, 20100101, 20100102, 201001~
## $ victim_last  <chr> "GARCIA", "MONTOYA", "SATTERFIELD", "MENDIOLA", "MULA", ~
## $ victim_first  <chr> "JUAN", "CAMERON", "VIVIANA", "CARLOS", "VIVIAN", "GERAL~
## $ victim_race   <chr> "Hispanic", "Hispanic", "White", "Hispanic", "White", "W~
## $ victim_age    <chr> "78", "17", "15", "32", "72", "91", "52", "52", "56", "4~
## $ victim_sex    <chr> "Male", "Male", "Female", "Male", "Female", "Female", "M~
## $ city         <chr> "Albuquerque", "Albuquerque", "Albuquerque", "Albuquerque", ~
## $ state        <chr> "NM", "NM", "NM", "NM", "NM", "NM", "NM", "NM", "NM", "N~
## $ lat          <dbl> 35.09579, 35.05681, 35.08609, 35.07849, 35.13036, 35.151~
```

```
## $ lon      <dbl> -106.5386, -106.7153, -106.6956, -106.5561, -106.5810, ~
## $ disposition <chr> "Closed without arrest", "Closed by arrest", "Closed wit~
```

5. Create monthly homicide counts for Baltimore

```
baltimore_monthly <-
  homicides_raw |>
  dplyr::filter(city == "Baltimore") |>
  # reported_date is an integer like 20150412; convert to Date
  dplyr::mutate(
    date = lubridate::ymd(as.character(reported_date)),
    month = lubridate::floor_date(date, unit = "month"),
    season = dplyr::if_else(
      lubridate::month(month) %in% c(11, 12, 1, 2, 3, 4),
      "Cold months (Nov-Apr)",
      "Warm months (May-Oct)"
    )
  ) |>
  dplyr::count(month, season, name = "homicides")

head(baltimore_monthly)

## # A tibble: 6 x 3
##   month     season       homicides
##   <date>    <chr>        <int>
## 1 2007-01-01 Cold months (Nov-Apr)     28
## 2 2007-02-01 Cold months (Nov-Apr)     17
## 3 2007-03-01 Cold months (Nov-Apr)     26
## 4 2007-04-01 Cold months (Nov-Apr)     19
## 5 2007-05-01 Warm months (May-Oct)    32
## 6 2007-06-01 Warm months (May-Oct)    31
```

6. Plot monthly homicides in Baltimore

```
freddie_arrest <- as.Date("2015-04-12")

ggplot2::ggplot(baltimore_monthly,
                ggplot2::aes(x = month, y = homicides, color = season)) +
  ggplot2::geom_line() +
  ggplot2::geom_point(size = 1.5) +
  # smooth line for seasonal / long-term trends
  ggplot2::geom_smooth(se = FALSE, size = 1, linetype = "solid") +
  ggplot2::geom_vline(xintercept = freddie_arrest, linetype = "dashed") +
  ggplot2::annotate(
    "text",
    x = freddie_arrest,
    y = max(baltimore_monthly$homicides, na.rm = TRUE) + 2,
    label = "Freddie Gray arrest\nApr 12, 2015",
    hjust = -0.05,
    vjust = 1,
    size = 3
  ) +
  ggplot2::scale_color_manual(
    values = c(
```

```

    "Cold months (Nov-Apr)" = "#1f78b4",
    "Warm months (May-Oct)" = "#e31a1c"
)
) +
ggplot2::scale_x_date(
  date_breaks = "1 year",
  date_labels = "%Y"
) +
ggplot2::labs(
  title = "Monthly homicides in Baltimore",
  subtitle = "Cold vs warm months, with reference line for the arrest of Freddie Gray",
  x = "Month",
  y = "Number of homicides",
  color = "Season"
) +
ggplot2::theme_minimal(base_size = 11) +
ggplot2::theme(
  legend.position = "top",
  panel.grid.minor = ggplot2::element_blank()
)

```

Monthly homicides in Baltimore

Cold vs warm months, with reference line for the arrest of Freddie Gray

