Homework 5

Danielle Scott

2023-11-28

list.files("data/")

## [1] "homicide-data (1).csv"

library(tidyverse)

## Warning: package 'tidyverse' was built under R version 4.2.3

## Warning: package 'ggplot2' was built under R version 4.2.3

## Warning: package 'tibble' was built under R version 4.2.3

## Warning: package 'readr' was built under R version 4.2.3

## Warning: package 'dplyr' was built under R version 4.2.3

## Warning: package 'stringr' was built under R version 4.2.3

## Warning: package 'forcats' was built under R version 4.2.3

## Warning: package 'lubridate' was built under R version 4.2.3

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.2 ✔ readr 2.1.4  
## ✔ forcats 1.0.0 ✔ stringr 1.5.0  
## ✔ ggplot2 3.4.2 ✔ tibble 3.2.1  
## ✔ lubridate 1.9.2 ✔ tidyr 1.3.0  
## ✔ purrr 1.0.1   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the ]8;;http://conflicted.r-lib.org/conflicted package]8;; to force all conflicts to become errors

library(lubridate)

homicide <- read\_csv("data/homicide-data (1).csv")

## Rows: 52179 Columns: 12  
## ── Column specification ────────────────────────────────────────────────────────  
## Delimiter: ","  
## chr (9): uid, victim\_last, victim\_first, victim\_race, victim\_age, victim\_sex...  
## dbl (3): reported\_date, lat, lon  
##   
## ℹ Use `spec()` to retrieve the full column specification for this data.  
## ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

head(homicide)

## # A tibble: 6 × 12  
## uid reported\_date victim\_last victim\_first victim\_race victim\_age victim\_sex  
## <chr> <dbl> <chr> <chr> <chr> <chr> <chr>   
## 1 Alb-… 20100504 GARCIA JUAN Hispanic 78 Male   
## 2 Alb-… 20100216 MONTOYA CAMERON Hispanic 17 Male   
## 3 Alb-… 20100601 SATTERFIELD VIVIANA White 15 Female   
## 4 Alb-… 20100101 MENDIOLA CARLOS Hispanic 32 Male   
## 5 Alb-… 20100102 MULA VIVIAN White 72 Female   
## 6 Alb-… 20100126 BOOK GERALDINE White 91 Female   
## # ℹ 5 more variables: city <chr>, state <chr>, lat <dbl>, lon <dbl>,  
## # disposition <chr>

baltimore <- homicide %>%  
 filter(city == "Baltimore")  
  
unique(baltimore$city) #confirm only Baltimore included in new dataframe

## [1] "Baltimore"

baltimore <- baltimore %>%  
 rename(date = reported\_date) %>%  
 mutate(date = ymd(date))  
  
class(baltimore$date)

## [1] "Date"

baltimore\_v2 <- baltimore %>%  
 mutate(month = month(date, label = T),  
 year = year(date)) %>%  
 group\_by(year, month) %>%  
 summarise(monthly\_homicide = n())

## `summarise()` has grouped output by 'year'. You can override using the  
## `.groups` argument.

baltimore\_v2

## # A tibble: 132 × 3  
## # Groups: year [11]  
## year month monthly\_homicide  
## <dbl> <ord> <int>  
## 1 2007 Jan 28  
## 2 2007 Feb 17  
## 3 2007 Mar 26  
## 4 2007 Apr 19  
## 5 2007 May 32  
## 6 2007 Jun 31  
## 7 2007 Jul 30  
## 8 2007 Aug 25  
## 9 2007 Sep 22  
## 10 2007 Oct 19  
## # ℹ 122 more rows