Homework 1 Solution

Aadarsha Gopala Reddy

September 8, 2022

Contents

1. Load Packages

In a chunk of code, load the tidyverse package and any other packages you will use in this document.

`install.packages()` installs a package, if it exists. `library()` loads the package.

```
library("tidyverse")
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6     v purrr   0.3.4
## v tibble  3.1.8     v dplyr   1.0.9
## v tidyr   1.2.0     v stringr 1.4.1
## v readr   2.1.2     v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

2. Import Data

Import protest_data.csv as an object in R. Print the first few rows of the dataset using the head() function.

```
protest_data <- read_csv("protest_data.csv")
```

```
## Rows: 21282 Columns: 11
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (7): event_date, event_type, sub_event_type, admin1, source, notes, loca...
## dbl (4): year, fatalities, latitude, longitude
##
## 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(protest_data)
```

```
## # A tibble: 6 x 11
##   event_date    year event~1 sub_e~2 admin1 source notes fatal~3 locat~4 latit~5
```

```
##   <chr>        <dbl> <chr>   <chr>   <chr>  <chr>  <chr>   <dbl> <chr>     <dbl>
## 1 12 August 2~  2022 Protes~ Peacef~ Calif~ KTVU ~ On 1~       0 Martin~    38.0
## 2 12 August 2~  2022 Protes~ Peacef~ Texas  News ~ On 1~       0 Midland    32.0
## 3 12 August 2~  2022 Protes~ Peacef~ Calif~ NewsC~ On 1~       0 Palm S~    33.8
## 4 12 August 2~  2022 Protes~ Peacef~ Georg~ 11 Al~ On 1~       0 Atlanta    33.8
## 5 12 August 2~  2022 Protes~ Peacef~ New Y~ 13WHA~ On 1~       0 Batavia    43.0
## 6 12 August 2~  2022 Protes~ Peacef~ Delaw~ Fox29  On 1~       0 Wilmin~    39.7
## # ... with 1 more variable: longitude <dbl>, and abbreviated variable names
## #   1: event_type, 2: sub_event_type, 3: fatalities, 4: location, 5: latitude
```

3. Filter Data

Subset the protest dataset to only the protests that occurred in the year 2022. Make sure to save your changes to the dataset.

```
protest_data <- filter(protest_data, year == 2022)
head(protest_data)
```

```
## # A tibble: 6 x 11
##   event_date    year event~1 sub_e~2 admin1 source notes fatal~3 locat~4 latit~5
##   <chr>        <dbl> <chr>   <chr>   <chr>  <chr>  <chr>   <dbl> <chr>     <dbl>
## 1 12 August 2~  2022 Protes~ Peacef~ Calif~ KTVU ~ On 1~       0 Martin~    38.0
## 2 12 August 2~  2022 Protes~ Peacef~ Texas  News ~ On 1~       0 Midland    32.0
## 3 12 August 2~  2022 Protes~ Peacef~ Calif~ NewsC~ On 1~       0 Palm S~    33.8
## 4 12 August 2~  2022 Protes~ Peacef~ Georg~ 11 Al~ On 1~       0 Atlanta    33.8
## 5 12 August 2~  2022 Protes~ Peacef~ New Y~ 13WHA~ On 1~       0 Batavia    43.0
## 6 12 August 2~  2022 Protes~ Peacef~ Delaw~ Fox29  On 1~       0 Wilmin~    39.7
## # ... with 1 more variable: longitude <dbl>, and abbreviated variable names
## #   1: event_type, 2: sub_event_type, 3: fatalities, 4: location, 5: latitude
```

```
protest_data %>% tally()
```

```
## # A tibble: 1 x 1
##       n
##   <int>
## 1  8305
```

How many protests occurred in the year 2022? **8305 protests**