Assignment 1

Mayuree

2022-06-04

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.1 ──

## ✓ ggplot2 3.3.5 ✓ purrr 0.3.4  
## ✓ tibble 3.1.6 ✓ dplyr 1.0.8  
## ✓ tidyr 1.2.0 ✓ stringr 1.4.0  
## ✓ readr 2.1.2 ✓ forcats 0.5.1

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

#import csv data

## New names:  
## \* `` -> ...1

## Rows: 100798 Columns: 11  
## ── Column specification ────────────────────────────────────────────────────────  
## Delimiter: ","  
## chr (6): month, intent, sex, race, place, education  
## dbl (5): ...1, year, police, age, hispanic  
##   
## 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.

## spec\_tbl\_df [100,798 × 11] (S3: spec\_tbl\_df/tbl\_df/tbl/data.frame)  
## $ ...1 : num [1:100798] 1 2 3 4 5 6 7 8 9 10 ...  
## $ year : num [1:100798] 2012 2012 2012 2012 2012 ...  
## $ month : chr [1:100798] "01" "01" "01" "02" ...  
## $ intent : chr [1:100798] "Suicide" "Suicide" "Suicide" "Suicide" ...  
## $ police : num [1:100798] 0 0 0 0 0 0 0 0 0 0 ...  
## $ sex : chr [1:100798] "M" "F" "M" "M" ...  
## $ age : num [1:100798] 34 21 60 64 31 17 48 41 50 NA ...  
## $ race : chr [1:100798] "Asian/Pacific Islander" "White" "White" "White" ...  
## $ hispanic : num [1:100798] 100 100 100 100 100 100 100 100 100 998 ...  
## $ place : chr [1:100798] "Home" "Street" "Other specified" "Home" ...  
## $ education: chr [1:100798] "BA+" "Some college" "BA+" "BA+" ...  
## - attr(\*, "spec")=  
## .. cols(  
## .. ...1 = col\_double(),  
## .. year = col\_double(),  
## .. month = col\_character(),  
## .. intent = col\_character(),  
## .. police = col\_double(),  
## .. sex = col\_character(),  
## .. age = col\_double(),  
## .. race = col\_character(),  
## .. hispanic = col\_double(),  
## .. place = col\_character(),  
## .. education = col\_character()  
## .. )  
## - attr(\*, "problems")=<externalptr>

#filter age less than or equal to 65

## spec\_tbl\_df [85,111 × 11] (S3: spec\_tbl\_df/tbl\_df/tbl/data.frame)  
## $ ...1 : num [1:85111] 1 2 3 4 5 6 7 8 9 11 ...  
## $ year : num [1:85111] 2012 2012 2012 2012 2012 ...  
## $ month : chr [1:85111] "01" "01" "01" "02" ...  
## $ intent : chr [1:85111] "Suicide" "Suicide" "Suicide" "Suicide" ...  
## $ police : num [1:85111] 0 0 0 0 0 0 0 0 0 0 ...  
## $ sex : chr [1:85111] "M" "F" "M" "M" ...  
## $ age : num [1:85111] 34 21 60 64 31 17 48 41 50 30 ...  
## $ race : chr [1:85111] "Asian/Pacific Islander" "White" "White" "White" ...  
## $ hispanic : num [1:85111] 100 100 100 100 100 100 100 100 100 100 ...  
## $ place : chr [1:85111] "Home" "Street" "Other specified" "Home" ...  
## $ education: chr [1:85111] "BA+" "Some college" "BA+" "BA+" ...  
## - attr(\*, "spec")=  
## .. cols(  
## .. ...1 = col\_double(),  
## .. year = col\_double(),  
## .. month = col\_character(),  
## .. intent = col\_character(),  
## .. police = col\_double(),  
## .. sex = col\_character(),  
## .. age = col\_double(),  
## .. race = col\_character(),  
## .. hispanic = col\_double(),  
## .. place = col\_character(),  
## .. education = col\_character()  
## .. )  
## - attr(\*, "problems")=<externalptr>

#Summarizing youth dataframe

## ...1 year month intent   
## Min. : 1 Min. :2012 Length:85111 Length:85111   
## 1st Qu.: 24944 1st Qu.:2012 Class :character Class :character   
## Median : 50035 Median :2013 Mode :character Mode :character   
## Mean : 50203 Mean :2013   
## 3rd Qu.: 75512 3rd Qu.:2014   
## Max. :100798 Max. :2014   
## police sex age race   
## Min. :0.00000 Length:85111 Min. : 0.00 Length:85111   
## 1st Qu.:0.00000 Class :character 1st Qu.:25.00 Class :character   
## Median :0.00000 Mode :character Median :36.00 Mode :character   
## Mean :0.01612 Mean :37.89   
## 3rd Qu.:0.00000 3rd Qu.:51.00   
## Max. :1.00000 Max. :65.00   
## hispanic place education   
## Min. :100.0 Length:85111 Length:85111   
## 1st Qu.:100.0 Class :character Class :character   
## Median :100.0 Mode :character Mode :character   
## Mean :115.6   
## 3rd Qu.:100.0   
## Max. :998.0

We have data about 100798 individuals killed by guns . Only 15687 are older than 65. The distribution of the remainder is shown below:



