AIT-Final Project

Naga Tejaswi Veluri

01/12/2021

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.0.5
```

```
##
## 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
```

```
data=read.csv("D:/study/Gmu/Sem 1/AIT/final/police_shootings.csv")
```

```
glimpse(data)
```

```
## Rows: 4,294
## Columns: 16
## $ Person.Name             <chr> "Tim Elliot", "Lewis Lee Lembke", "John Pa
ul Q~
## $ Person.Age              <int> 53, 47, 23, 32, 39, 18, 22, 35, 34, 47, 25
, 31~
## $ Person.Gender           <chr> "Male", "Male", "Male", "Male", "Male", "M
ale"~
## $ Person.Race             <chr> "Asian", "White", "Hispanic", "White", "Hi
span~
## $ Incident.Date.Month     <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1~
## $ Incident.Date.Day       <int> 2, 2, 3, 4, 4, 4, 5, 6, 6, 6, 6, 7, 7, 7,
7, 8~
## $ Incident.Date.Year      <int> 2015, 2015, 2015, 2015, 2015, 2015, 2015,
2015~
## $ Incident.Date.Full      <chr> "2015/01/02", "2015/01/02", "2015/01/03",
"201~
## $ Incident.Location.City  <chr> "Shelton", "Aloha", "Wichita", "San Franci
sco"~
## $ Incident.Location.State <chr> "WA", "OR", "KS", "CA", "CO", "OK", "AZ",
"KS"~
## $ Factors.Armed           <chr> "gun", "gun", "unarmed", "toy weapon", "na
il g~
```

```
## $ Factors.Mental.Illness  <chr> "True", "True", "True", "True", "True", "T
rue"~
## $ Factors.Threat.Level    <chr> "attack", "attack", "other", "attack", "at
tack~
## $ Factors.Fleeing         <chr> "Not fleeing", "Not fleeing", "Not fleeing
", "~
## $ Shooting.Manner         <chr> "shot", "shot", "shot and Tasered", "shot"
, "s~
## $ Shooting.Body.Camera    <chr> "True", "True", "True", "True", "True", "T
rue"~
```

This data set does not contain any null values

```
is.null(data)

## [1] FALSE
```

Replace 0 with mean

```
data[data == 0] <- round(mean(data$Person.Age))                # Replace 0 w
ith mean
summary(data)

##  Person.Name          Person.Age    Person.Gender      Person.Race
##  Length:4294        Min.   : 6.00   Length:4294        Length:4294
##  Class :character   1st Qu.:27.00   Class :character   Class :character
##  Mode  :character   Median :35.00   Mode  :character   Mode  :character
##                     Mean   :36.76
##                     3rd Qu.:45.00
##                     Max.   :91.00
##  Incident.Date.Month Incident.Date.Day Incident.Date.Year Incident.Date.Fu
ll
##  Min.   : 1.000      Min.   : 1.00     Min.   :2015       Length:4294
##  1st Qu.: 3.000      1st Qu.: 8.00     1st Qu.:2016       Class :character
##  Median : 6.000      Median :15.00     Median :2017       Mode  :character
##  Mean   : 6.061      Mean   :15.53     Mean   :2017
##  3rd Qu.: 9.000      3rd Qu.:23.00     3rd Qu.:2018
##  Max.   :12.000      Max.   :31.00     Max.   :2019
##  Incident.Location.City Incident.Location.State Factors.Armed
##  Length:4294            Length:4294             Length:4294
##  Class :character       Class :character        Class :character
##  Mode  :character       Mode  :character        Mode  :character
##
##
##
##  Factors.Mental.Illness Factors.Threat.Level Factors.Fleeing
##  Length:4294            Length:4294          Length:4294
##  Class :character       Class :character     Class :character
##  Mode  :character       Mode  :character     Mode  :character
```

```
##
##
##
##  Shooting.Manner    Shooting.Body.Camera
##  Length:4294        Length:4294
##  Class :character   Class :character
##  Mode  :character   Mode  :character
##
##
##

data1=data.frame(data)
head(data1)

##          Person.Name Person.Age Person.Gender Person.Race Incident.Date.Mo
nth
## 1         Tim Elliot         53          Male       Asian
1
## 2   Lewis Lee Lembke         47          Male       White
1
## 3 John Paul Quintero         23          Male    Hispanic
1
## 4    Matthew Hoffman         32          Male       White
1
## 5  Michael Rodriguez         39          Male    Hispanic
1
## 6  Kenneth Joe Brown         18          Male       White
1
##   Incident.Date.Day Incident.Date.Year Incident.Date.Full
## 1                 2               2015         2015/01/02
## 2                 2               2015         2015/01/02
## 3                 3               2015         2015/01/03
## 4                 4               2015         2015/01/04
## 5                 4               2015         2015/01/04
## 6                 4               2015         2015/01/04
##   Incident.Location.City Incident.Location.State Factors.Armed
## 1               Shelton                      WA           gun
## 2                 Aloha                      OR           gun
## 3               Wichita                      KS       unarmed
## 4         San Francisco                      CA    toy weapon
## 5                 Evans                      CO      nail gun
## 6               Guthrie                      OK           gun
##   Factors.Mental.Illness Factors.Threat.Level Factors.Fleeing  Shooting.Ma
nner
## 1                   True               attack     Not fleeing
shot
## 2                   True               attack     Not fleeing
shot
## 3                   True                other     Not fleeing shot and Tas
ered
```

```
## 4                   True               attack     Not fleeing
shot
## 5                   True               attack     Not fleeing
shot
## 6                   True               attack     Not fleeing
shot
##   Shooting.Body.Camera
## 1                 True
## 2                 True
## 3                 True
## 4                 True
## 5                 True
## 6                 True
```

Finding Numeric Data types

```
continuous=select_if(data1,is.numeric)#just taking numeric vlues
summary(continuous)
```

```
##    Person.Age    Incident.Date.Month Incident.Date.Day Incident.Date.Year
##  Min.   : 6.00   Min.   : 1.000      Min.   : 1.00     Min.   :2015
##  1st Qu.:27.00   1st Qu.: 3.000      1st Qu.: 8.00     1st Qu.:2016
##  Median :35.00   Median : 6.000      Median :15.00     Median :2017
##  Mean   :36.76   Mean   : 6.061      Mean   :15.53     Mean   :2017
##  3rd Qu.:45.00   3rd Qu.: 9.000      3rd Qu.:23.00     3rd Qu.:2018
##  Max.   :91.00   Max.   :12.000      Max.   :31.00     Max.   :2019
```

Renaming the columns in dataset

```
names(data1) <- c('Name','Age','Gender','Race','Month','Day','Year','Incident
Date','City','State','Armed','Mental_Illness','ThreatLevel','Fleeing','Manner
','BodyCamera')

head(data1)
```

```
##                 Name Age Gender     Race Month Day Year IncidentDate
## 1         Tim Elliot  53   Male    Asian     1   2 2015   2015/01/02
## 2   Lewis Lee Lembke  47   Male    White     1   2 2015   2015/01/02
## 3 John Paul Quintero  23   Male Hispanic     1   3 2015   2015/01/03
## 4    Matthew Hoffman  32   Male    White     1   4 2015   2015/01/04
## 5  Michael Rodriguez  39   Male Hispanic     1   4 2015   2015/01/04
## 6  Kenneth Joe Brown  18   Male    White     1   4 2015   2015/01/04
##            City State      Armed Mental_Illness ThreatLevel     Fleeing
## 1       Shelton    WA        gun           True      attack Not fleeing
## 2         Aloha    OR        gun           True      attack Not fleeing
## 3       Wichita    KS    unarmed           True       other Not fleeing
## 4 San Francisco    CA toy weapon           True      attack Not fleeing
## 5         Evans    CO   nail gun           True      attack Not fleeing
## 6       Guthrie    OK        gun           True      attack Not fleeing
```

```
##              Manner BodyCamera
## 1              shot       True
## 2              shot       True
## 3 shot and Tasered       True
## 4              shot       True
## 5              shot       True
## 6              shot       True
```

Saving the data into CSV file

```
write.csv(data1,"D:/study/Gmu/Sem 1/AIT/final/police_shootings_modifiedset.csv",row.names = TRUE )
```