

# GEO\_Code - National Highway Traffic Safety Administration (NHTSA)

J. Dayton

5/13/2020

## National Highway Traffic Safety Administration (NHTSA) - Fatality Analysis Reporting System (FARS)

### STEP 0: Load Libraries

#### Get the Vehicle Crash Data

Source of data: <https://www.nhtsa.gov/node/97996/176776>.

National Highway Traffic Safety Administration (NHTSA) - Fatality Analysis Reporting System (FARS)

#### FARS Auxiliary Files

The FARS Auxiliary files are datasets that are derived from the standard FARS datasets:

- 1. Accident -> Acc\_Aux
- 2. Vehicle -> Veh\_Aux
- 3. Person -> Per\_Aux

These files are joined by the standard key variables. Please read the following document for additional details: <http://www-nrd.nhtsa.dot.gov/Pubs/811364.pdf> For User Manuals, please go to: [ftp://ftp.nhtsa.dot.gov/FARS/FARS-Doc/FARS\\_Auxiliary\\_Analytical\\_User\\_Manuals.zip](ftp://ftp.nhtsa.dot.gov/FARS/FARS-Doc/FARS_Auxiliary_Analytical_User_Manuals.zip) For SAS Format Catalog, please go to: [ftp://ftp.nhtsa.dot.gov/FARS/Auxiliary\\_FARS\\_Files\\_Formats](ftp://ftp.nhtsa.dot.gov/FARS/Auxiliary_FARS_Files_Formats)

### STEP 1: Import relevant data set.

```
##Load the data
dfAcc <- read.csv("./FARS2018NationalCSV/ACCIDENT.csv",
                  stringsAsFactors = FALSE)
```

#### Clean Data

Check for NA values. Due to length of output, code commented out to not show in this document.

```
# head(dfAcc)
# summary(dfAcc)
dim(dfAcc)
```

```
## [1] 33654    52
```

## STEP 2. Aggregate data.

Focus data on area of interest:

- STATE = 10 for Delaware
- STATE = 11 for Washington, DC
- STATE = 24 for Maryland
- STATE = 51 for Virginia

```
df <- dfAcc[dfAcc$STATE %in% c(10, 11, 24, 51), ] #For DE, DC, MD, VA
# rm(dfAcc) #Removing original df - no longer referenced
dim(df)
```

```
## [1] 1386    52
```

Fix the date by combining multiple elements from original data.

```
df$Date <- mdy_hms(paste0(df$MONTH, "/", df$DAY, "/", df$YEAR, " ", df$HOUR, ":", df$MINUTE, ":00"))
```

Reduce the data frame to focus on the items of interest. Also, renaming the variables to simplify.

```
df <- df %>%
  select(Date = Date, Lat = LATITUDE, Lon = LONGITUD, State = STATE,
         Pedestrians = PERNOTMVIT, Motorists = PERMVIT, Fatalities = FATALS,
         Vehicles = VE_TOTAL)
dim(df)
```

```
## [1] 1386     8
```

```
kable(head(df), "pandoc", row.names = FALSE)
```

Date	Lat	Lon	State	Pedestrians	Motorists	Fatalities	Vehicles
2018-03-16 19:00:00	38.90829	-75.6582	10	0	1	1	1
2018-01-06 02:09:00	39.65505	-75.7025	10	0	1	1	1
2018-01-08 11:09:00	38.51757	-75.2400	10	0	5	1	2
2018-01-10 04:05:00	39.55601	-75.7033	10	0	1	1	1
2018-01-10 07:49:00	39.72868	-75.5382	10	0	2	1	2
2018-01-22 15:39:00	39.60879	-75.6642	10	0	2	1	2

### STEP 3. Merge data from data set to map data.

Not required for leaflet plot when data contains lat & lon.

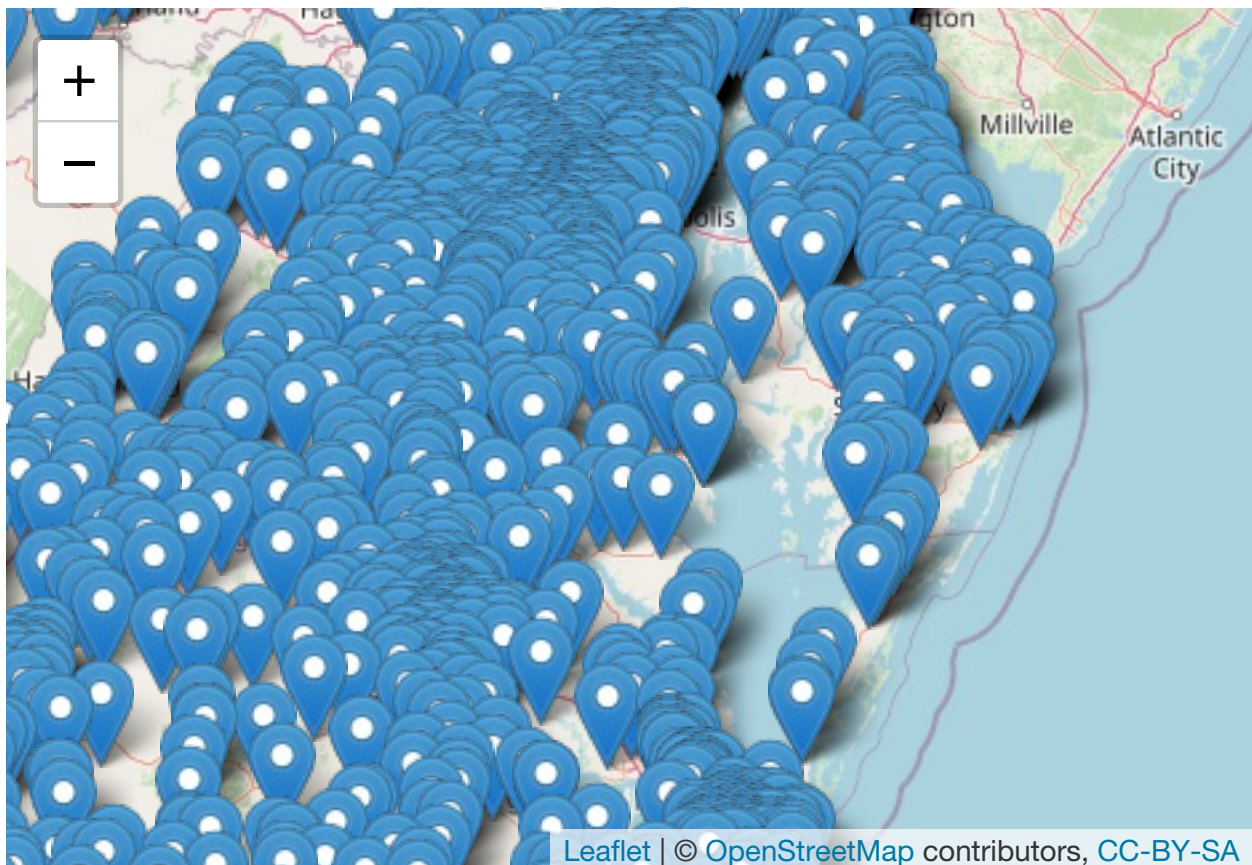
### STEP 4. Create the plot(s).

#### Leaflet with Markers

```
labs <- lapply(seq(nrow(df)), function(i) {  
  paste0( '

<b>DATE: </b>', format(df[i, "Date"], format="%d %b %y"),  
    '</p><p><b>TIME: </b>', format(df[i, "Date"], format = "%H:%M"),  
    '</p><p><b>FATALITIES: </b>',df[i, "Fatalities"],  
    '</p><p><b>VEHICLES: </b>',df[i, "Vehicles"],  
    '</p><p><b>MOTORISTS: </b>',df[i, "Motorists"],  
    '</p><p><b>PEDESTRIANS: </b>',df[i, "Pedestrians"], '</p>' )  
}  
)  
  
leafMap <- df %>%  
  leaflet() %>%  
  addTiles() %>%  
  setView(lng = mean(df$Lon, na.rm = TRUE), lat = mean(df$Lat, na.rm = TRUE), zoom = 7) %>%  
  addMarkers(lng = ~Lon, lat = ~Lat,  
    label = lapply(labs, htmltools::HTML) )  
leafMap

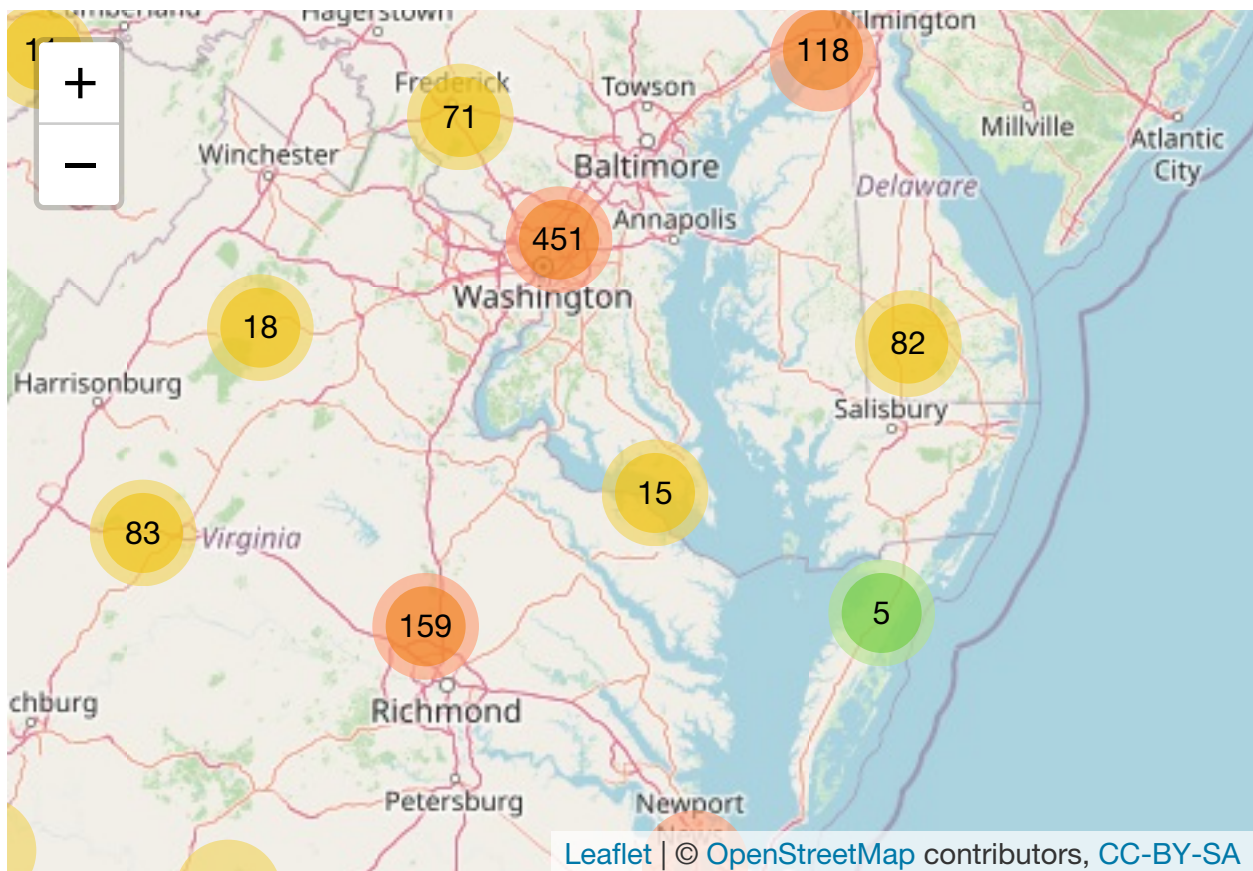

```



## 2018 Motor Vehicle Fatalities in the DMV (above)

### Leaflet with Cluster Markers

```
leafClustMap <- df %>%  
  leaflet() %>%  
  addTiles() %>%  
  setView(lng = mean(df$Lon, na.rm = TRUE), lat = mean(df$Lat, na.rm = TRUE), zoom = 7) %>%  
  addMarkers(lng = ~Lon, lat = ~Lat,  
             label = lapply(labs, htmltools::HTML),  
             clusterOptions = markerClusterOptions() )  
leafClustMap
```



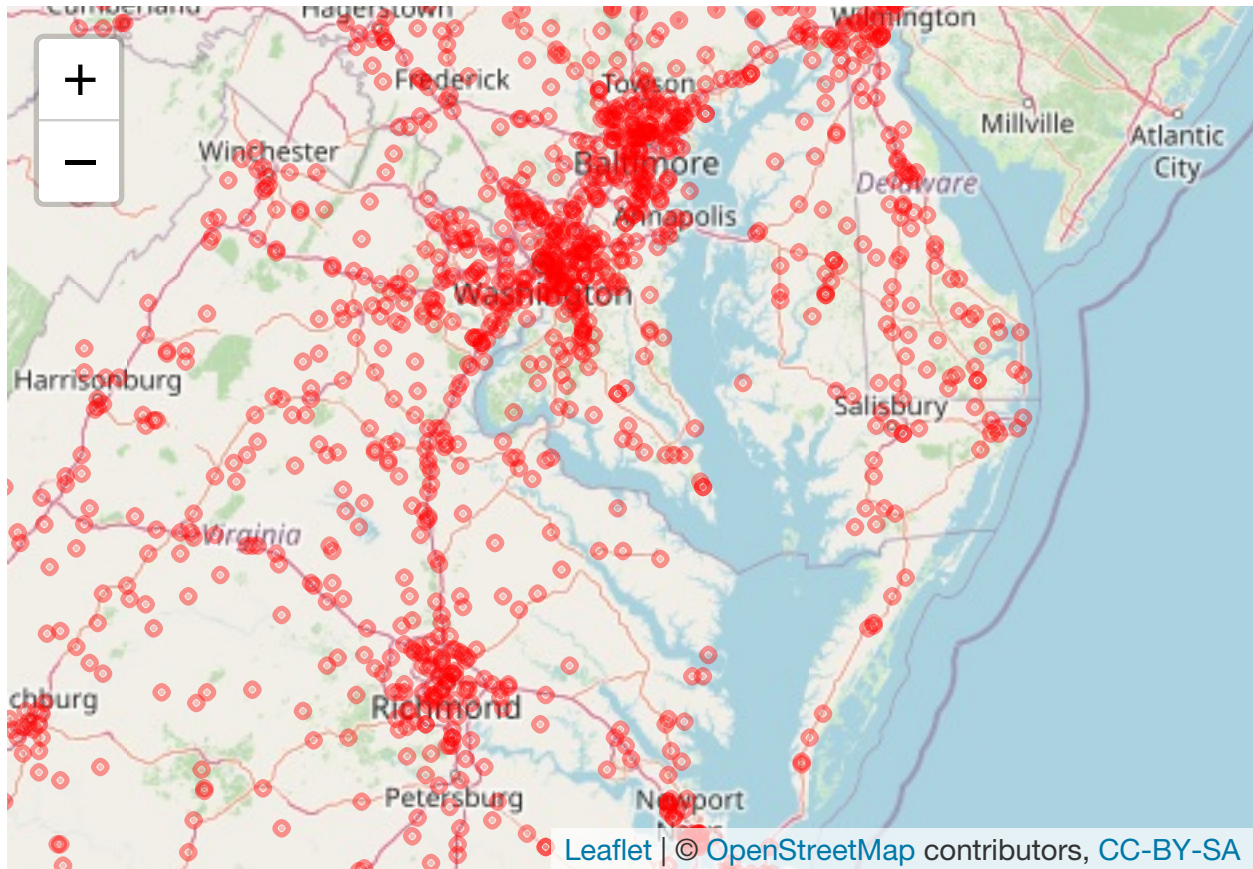
2018 Motor Vehicle Fatalities in the DMV (above)

### Leaflet with Circle Markers

```
leafCircleMap <- df %>%  
  leaflet() %>%  
  addTiles() %>%  
  setView(lng = mean(df$Lon, na.rm = TRUE), lat = mean(df$Lat, na.rm = TRUE), zoom = 7) %>%  
  addCircleMarkers(lng = ~Lon, lat = ~Lat,
```



```
color = "red", opacity = 0.4, radius = 0.5,  
label = lapply(labs, htmltools::HTML) )  
leafCircleMap
```



2018 Motor Vehicle Fatalities in the DMV (above)