Reproducible Research Project 2

lnma

4/27/2020

1.Synopsis

Storms and other severe weather events can cause both public health and economic problems for communities and municipalities. Many severe events can result in fatalities, injuries, and property damage, and preventing such outcomes to the extent possible is a key concern.
This project involves exploring the U.S. National Oceanic and Atmospheric Administration's (NOAA) storm database. This database tracks characteristics of major storms and weather events in the United States, including when and where they occur, as well as estimates of any fatalities, injuries, and property damage.
In this report,effect of weather events on damage of people and property was studied. Barplots of the top 10 weather events that cause highest fatalities and injuries were plottet, which show that the most fatalities and injuries were caused by *Tornados*. In addition, barplot of the top 10 weather events resulting in the highest total economic damage shows that *Flood* caused the most economic damage.

2.Data download and process

2.1 Data

The data for this assignment come in the form of a comma-separated-value file compressed via the bzip2 algorithm to reduce its size. You can download the file from the course web site:

- Storm Data

2.2 Data download and load

The data was downloaded and read the variables of *EVTYPE*, *FATALITIES*, *INJURIES*, *PROPDMG*, *PROPDMGEXP*, *CROPDMG*, *CROPDMGEXP* into a file with name *event*

```
url <- "https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2"
destfile <- "data.csv"
download.file(url, destfile)
data <- read.csv("data.csv", sep = ",")
event<-data[,c("EVTYPE","FATALITIES","INJURIES","PROPDMG","PROPDMGEXP","CROPDMG","CROPDMGEXP")]
```

2.3 Data process

```
head(event)
```

2.3.1 Look at the raw data

```
##    EVTYPE FATALITIES INJURIES PROPDMG PROPDMGEXP CROPDMG CROPDMGEXP
## 1 TORNADO          0       15    25.0          K       0
## 2 TORNADO          0        0     2.5          K       0
## 3 TORNADO          0        2    25.0          K       0
## 4 TORNADO          0        2     2.5          K       0
```

```
## 5 TORNADO          0        2     2.5          K       0
## 6 TORNADO          0        6     2.5          K       0
```

2.3.2 Health impact To evaluate the health impact, the total fatalities and the total injuries for each event type (EVTYPE) are calculated. The codes for this calculation are shown as follows.

```
library(reshape2)
fat_injur<-melt(event,id='EVTYPE',measure.vars = c('INJURIES','FATALITIES'))
fat_injur_sum<-dcast(fat_injur,EVTYPE~variable,sum)
head(fat_injur_sum)
```

```
##                  EVTYPE INJURIES FATALITIES
## 1    HIGH SURF ADVISORY        0          0
## 2         COASTAL FLOOD        0          0
## 3           FLASH FLOOD        0          0
## 4             LIGHTNING        0          0
## 5             TSTM WIND        0          0
## 6       TSTM WIND (G45)        0          0
```

The Top 10 Events with highest health impact are calcualted. The codes for this calculation are shown as follows.

```
fat_sum<-fat_injur_sum[order(-fat_injur_sum$FATALITIES),][,c('EVTYPE','FATALITIES')]
injur_sum<-fat_injur_sum[order(-fat_injur_sum$INJURIES),][,c('EVTYPE','INJURIES')]
head(fat_sum,10)
```

```
##             EVTYPE FATALITIES
## 834        TORNADO       5633
## 130 EXCESSIVE HEAT       1903
## 153    FLASH FLOOD        978
## 275           HEAT        937
## 464      LIGHTNING        816
## 856      TSTM WIND        504
## 170          FLOOD        470
## 585    RIP CURRENT        368
## 359      HIGH WIND        248
## 19       AVALANCHE        224
```

```
head(injur_sum,10)
```

```
##               EVTYPE INJURIES
## 834          TORNADO    91346
## 856        TSTM WIND     6957
## 170            FLOOD     6789
## 130   EXCESSIVE HEAT     6525
## 464        LIGHTNING     5230
## 275             HEAT     2100
## 427        ICE STORM     1975
## 153      FLASH FLOOD     1777
## 760 THUNDERSTORM WIND     1488
## 244             HAIL     1361
```

2.3.3 Economic impact The data provides two types of economic impact, namely property damage (PROPDMG) and crop damage (CROPDMG). The actual damage in *USD* is indicated by PROPDMGEXP and CROPDMGEXP parameters. According to the instructions, the characters in the PROPDMGEXP and CROPDMGEXP can be interpreted as the following:

- H,h = hundreds = 100
- k,k = kilos = thousands = 1,000
- M,m = millions = 1,000,000
- B,b = billions = 1,000,000,000
- (+) = 1
- (-) = 0
- (?) = 0
- black/empty character = 0
- numeric 0..8 = 10

The damage caused by each event type is calculated with the following code.

```
library(dplyr)
damage<-event%>%select(EVTYPE,PROPDMG,PROPDMGEXP,CROPDMG,CROPDMGEXP)
symbol<-sort(unique(as.character(damage$PROPDMGEXP)))
number<-c(0,0,0,1,10,10,10,10,10,10,10,10,10,1e9,1e2,1e2,1e3,1e6,1e6)
damage$PROPDMGEXP<-number[match(damage$PROPDMGEXP,symbol)]
damage$CROPDMGEXP<-number[match(damage$CROPDMGEXP,symbol)]
damage<-damage%>%mutate(PROPDMGtotal=PROPDMG*PROPDMGEXP)%>%
        mutate(CROPDMGtotal=CROPDMG*CROPDMGEXP)%>%
        mutate(DamageTotal=PROPDMGtotal+CROPDMGtotal)
head(damage)
```

```
##    EVTYPE PROPDMG PROPDMGEXP CROPDMG CROPDMGEXP PROPDMGtotal CROPDMGtotal
## 1 TORNADO    25.0       1000       0          0        25000            0
## 2 TORNADO     2.5       1000       0          0         2500            0
## 3 TORNADO    25.0       1000       0          0        25000            0
## 4 TORNADO     2.5       1000       0          0         2500            0
## 5 TORNADO     2.5       1000       0          0         2500            0
## 6 TORNADO     2.5       1000       0          0         2500            0
##   DamageTotal
## 1       25000
## 2        2500
## 3       25000
## 4        2500
## 5        2500
## 6        2500
```

The total damage caused by each event type is calculated and ordered from highest to lowest with the following code. The number was reshaped to billion by dividing the variable by *1e9*.

```
DamageTotal<-damage%>%group_by(EVTYPE)%>%
        summarize(TotalDamage=sum(DamageTotal))%>%
        arrange(-TotalDamage)
DamageTotal$TotalDamage=DamageTotal$TotalDamage/1e9 #to billion
head(DamageTotal)
```

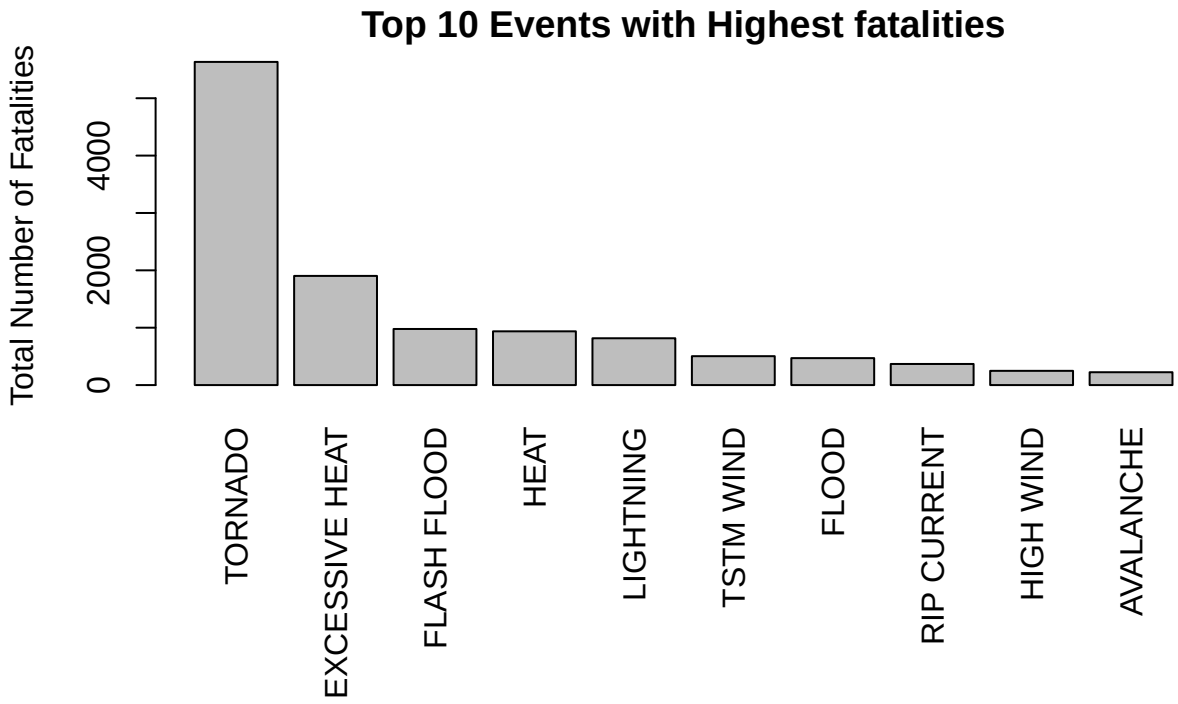
```
## # A tibble: 6 x 2
##   EVTYPE            TotalDamage
##   <fct>                   <dbl>
## 1 FLOOD                   150.
## 2 HURRICANE/TYPHOON        71.9
## 3 TORNADO                  57.4
## 4 STORM SURGE              43.3
## 5 FLASH FLOOD              17.6
## 6 DROUGHT                  15.0
```

3.Results

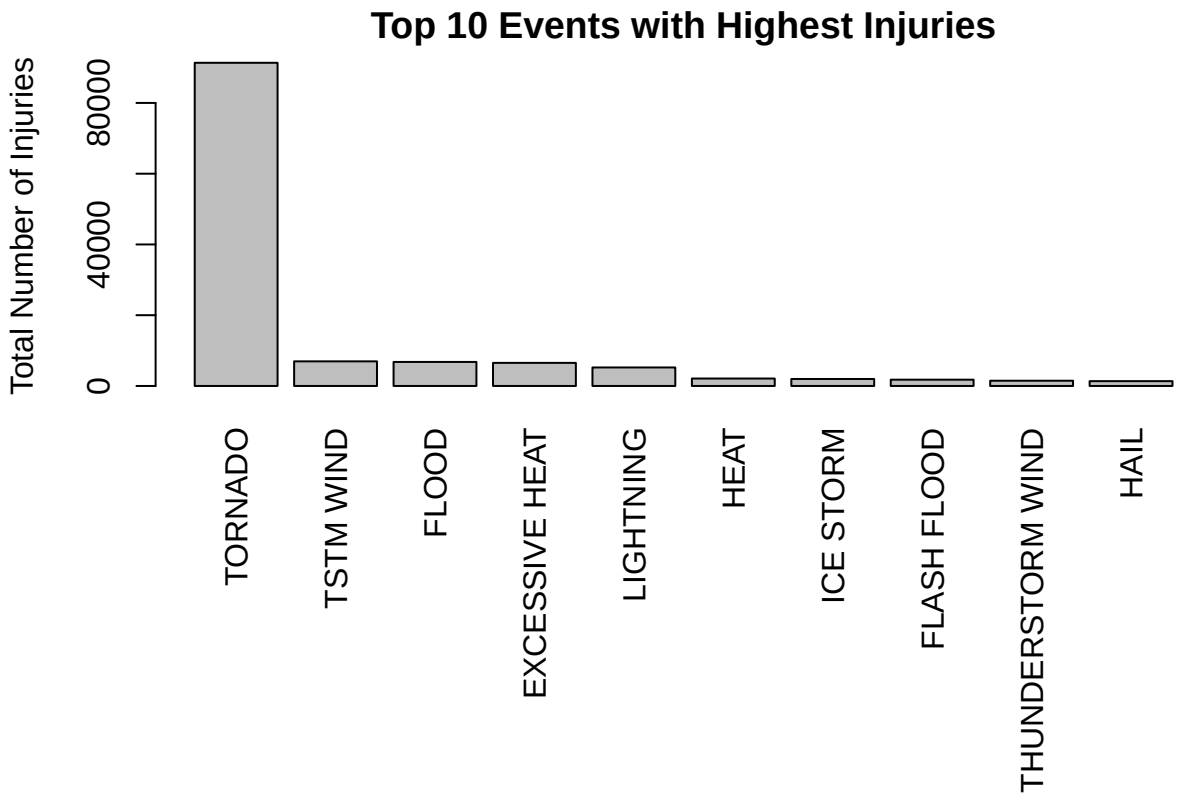
3.1 Health impact

The top 10 events with the highest total fatalities and injuries are shown graphically.

```
par(mfrow=c(1,1),mar=c(12,4,2,1))
barplot(fat_sum$FATALITIES[1:10],names.arg = fat_sum$EVTYPE[1:10],las=3,
        main = 'Top 10 Events with Highest fatalities',
        ylab = 'Total Number of Fatalities',)
```




```
par(mfrow=c(1,1),mar=c(12,4,2,1))
barplot(injur_sum$INJURIES[1:10],names.arg = injur_sum$EVTYPE[1:10],las=3,
        main = 'Top 10 Events with Highest Injuries',
        ylab = 'Total Number of Injuries'
        )
```

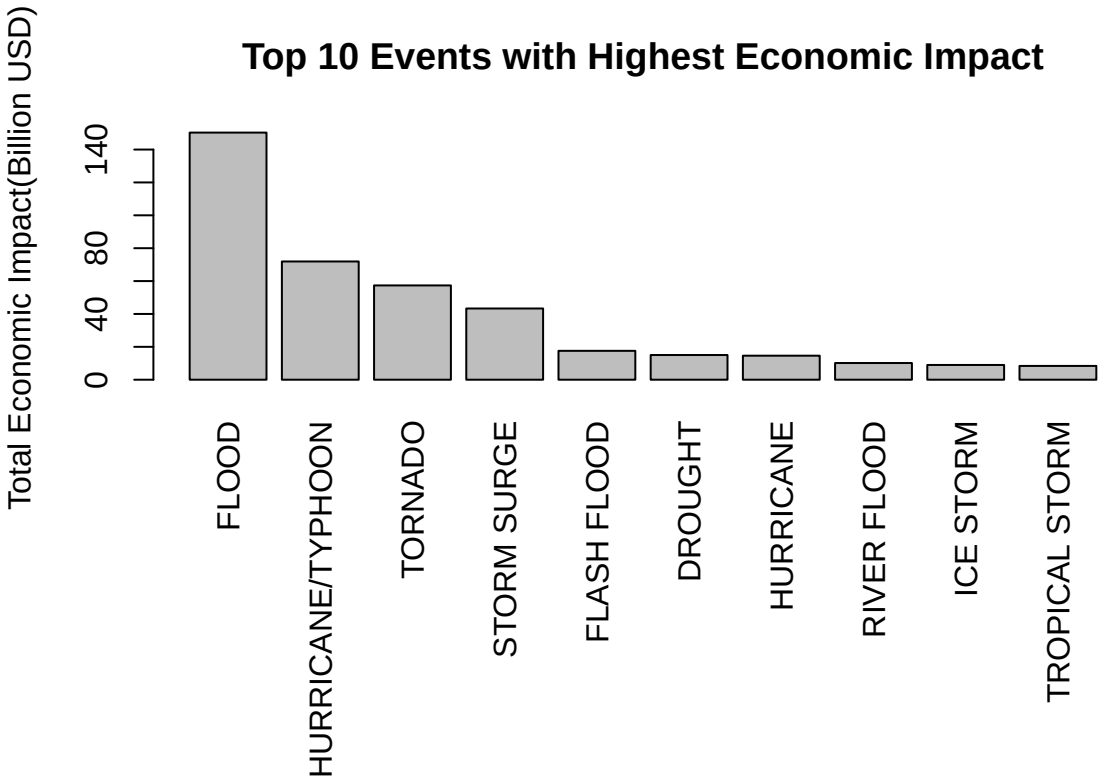


As shown in the figures, tornado causes the hightest in both the total fatalities and injuries.

3.2 Economic impact

The top 10 events with the highest total economic damages (property and crop combined) are shown graphically.

```
par(mfrow=c(1,1),mar=c(12,6,4,1))
barplot(DamageTotal$TotalDamage[1:10],names.arg = DamageTotal$EVTYPE[1:10],las=3,
        main = 'Top 10 Events with Highest Economic Impact',
        ylab = 'Total Economic Impact(Billion USD)')
```



As shown in the figures, flood causes the hightest economic damage. And the total loss is as high as 150 billion USD.

4.Conclusions

Tornados caused the maximum number of fatalities and injuries. It was followed by Excessive Heat for fatalities and Thunderstorm wind for injuries.

Floods caused the maximum total economic damage which is followed by Hurricane/Typhoon, Tornado, Storm surge, etc.