Assgmnt-R

Abhishek R

24/10/2021

```
# reading data
Raw_data <- read.csv(file = "repdata_data_StormData.csv", header=TRUE, sep=",")

# subsetting by date
Main_data <- Raw_data
Main_data$BGN_DATE <- strptime(Raw_data$BGN_DATE, "%m/%d/%Y %H:%M:%S")
Main_data <- subset(Main_data, BGN_DATE > "1995-12-31")

# subsetting to needed columns
Main_data <- subset(Main_data, select = c(EVTYPE, FATALITIES, INJURIES, PROPDMG, PROPDMGEXP, CROPDMG, C

#cleaning event types names
Main_data$EVTYPE <- toupper(Main_data$EVTYPE)

# eliminating zero data
Main_data <- Main_data[Main_data$FATALITIES !=0 |
                         Main_data$INJURIES !=0 |
                         Main_data$PROPDMG !=0 |
                         Main_data$CROPDMG !=0, ]

#---------Population health data processing

Health_data <- aggregate(cbind(FATALITIES, INJURIES) ~ EVTYPE, data = Main_data, FUN=sum)
Health_data$PEOPLE_LOSS <- Health_data$FATALITIES + Health_data$INJURIES
Health_data <- Health_data[order(Health_data$PEOPLE_LOSS, decreasing = TRUE), ]
Top10_events_people <- Health_data[1:10,]
print(Top10_events_people)
```

```
##                EVTYPE FATALITIES INJURIES PEOPLE_LOSS
## 149           TORNADO       1511    20667       22178
## 39     EXCESSIVE HEAT       1797     6391        8188
## 48              FLOOD        414     6758        7172
## 107         LIGHTNING        651     4141        4792
## 153         TSTM WIND        241     3629        3870
## 46        FLASH FLOOD        887     1674        2561
## 146 THUNDERSTORM WIND        130     1400        1530
## 182      WINTER STORM        191     1292        1483
## 69               HEAT        237     1222        1459
## 88  HURRICANE/TYPHOON         64     1275        1339
```

```
#---------Economic consequences data processing

#transforming letters and symbols to numbers

Main_data$PROPDMGEXP <- gsub("[Hh]", "2", Main_data$PROPDMGEXP)
Main_data$PROPDMGEXP <- gsub("[Kk]", "3", Main_data$PROPDMGEXP)
Main_data$PROPDMGEXP <- gsub("[Mm]", "6", Main_data$PROPDMGEXP)
Main_data$PROPDMGEXP <- gsub("[Bb]", "9", Main_data$PROPDMGEXP)
Main_data$PROPDMGEXP <- gsub("\\+", "1", Main_data$PROPDMGEXP)
Main_data$PROPDMGEXP <- gsub("\\?|\\-|\\ ", "0",  Main_data$PROPDMGEXP)
Main_data$PROPDMGEXP <- as.numeric(Main_data$PROPDMGEXP)

Main_data$CROPDMGEXP <- gsub("[Hh]", "2", Main_data$CROPDMGEXP)
Main_data$CROPDMGEXP <- gsub("[Kk]", "3", Main_data$CROPDMGEXP)
Main_data$CROPDMGEXP <- gsub("[Mm]", "6", Main_data$CROPDMGEXP)
Main_data$CROPDMGEXP <- gsub("[Bb]", "9", Main_data$CROPDMGEXP)
Main_data$CROPDMGEXP <- gsub("\\+", "1", Main_data$CROPDMGEXP)
Main_data$CROPDMGEXP <- gsub("\\-|\\?|\\ ", "0", Main_data$CROPDMGEXP)
Main_data$CROPDMGEXP <- as.numeric(Main_data$CROPDMGEXP)

Main_data$PROPDMGEXP[is.na(Main_data$PROPDMGEXP)] <- 0
Main_data$CROPDMGEXP[is.na(Main_data$CROPDMGEXP)] <- 0

#creating total damage values
library(dplyr)
```

```
##
## 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
```

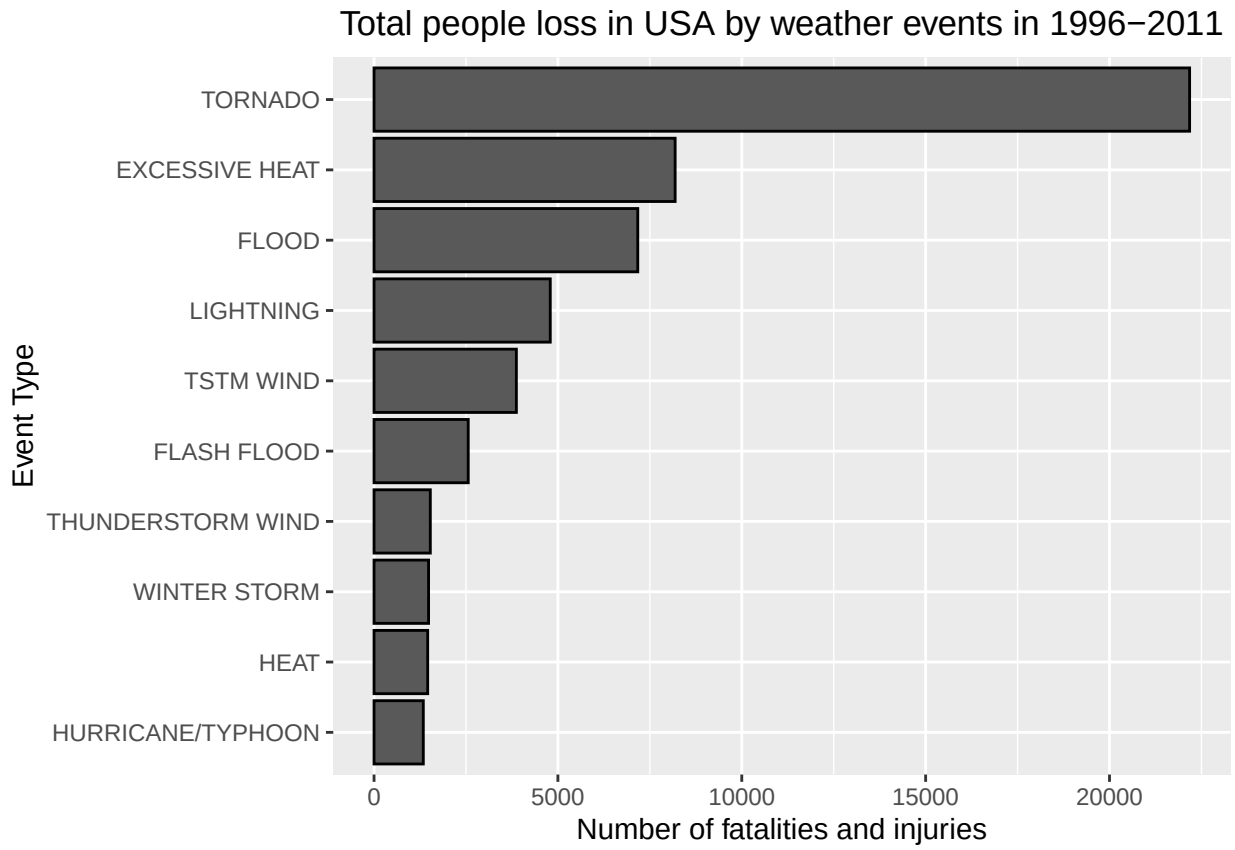
```
Main_data <- mutate(Main_data,
                    PROPDMGTOTAL = PROPDMG * (10 ^ PROPDMGEXP),
                    CROPDMGTOTAL = CROPDMG * (10 ^ CROPDMGEXP))

#analyzing
Economic_data <- aggregate(cbind(PROPDMGTOTAL, CROPDMGTOTAL) ~ EVTYPE, data = Main_data, FUN=sum)
Economic_data$ECONOMIC_LOSS <- Economic_data$PROPDMGTOTAL + Economic_data$CROPDMGTOTAL
Economic_data <- Economic_data[order(Economic_data$ECONOMIC_LOSS, decreasing = TRUE), ]
Top10_events_economy <- Economic_data[1:10,]
print(Top10_events_economy)
```

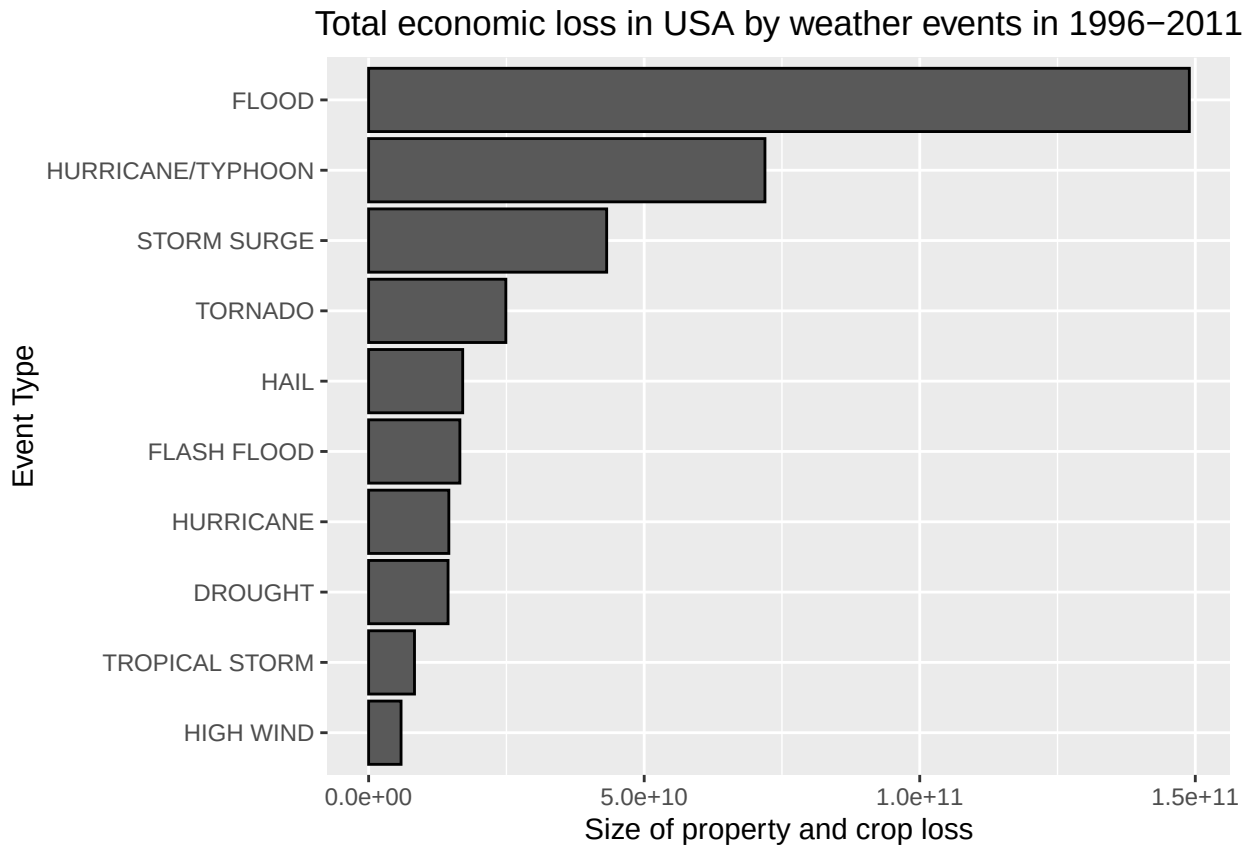
```
##                EVTYPE PROPDMGTOTAL CROPDMGTOTAL ECONOMIC_LOSS
## 48              FLOOD 143944833550   4974778400  148919611950
## 88  HURRICANE/TYPHOON  69305840000   2607872800   71913712800
## 141       STORM SURGE  43193536000         5000   43193541000
```

```
## 149           TORNADO  24616945710    283425010   24900370720
## 66               HAIL  14595143420   2476029450   17071172870
## 46        FLASH FLOOD  15222203910   1334901700   16557105610
## 86          HURRICANE  11812819010   2741410000   14554229010
## 32            DROUGHT   1046101000  13367566000   14413667000
## 152    TROPICAL STORM   7642475550    677711000    8320186550
## 83           HIGH WIND   5247860360    633561300    5881421660
```

```
#plotting health loss
library(ggplot2)
g <- ggplot(data = Top10_events_people, aes(x = reorder(EVTYPE, PEOPLE_LOSS), y = PEOPLE_LOSS))
g <- g + geom_bar(stat = "identity", colour = "black")
g <- g + labs(title = "Total people loss in USA by weather events in 1996-2011")
g <- g + theme(plot.title = element_text(hjust = 0.5))
g <- g + labs(y = "Number of fatalities and injuries", x = "Event Type")
g <- g + coord_flip()
print(g)
```




```
#plotting economic loss
g <- ggplot(data = Top10_events_economy, aes(x = reorder(EVTYPE, ECONOMIC_LOSS), y = ECONOMIC_LOSS))
g <- g + geom_bar(stat = "identity", colour = "black")
g <- g + labs(title = "Total economic loss in USA by weather events in 1996-2011")
g <- g + theme(plot.title = element_text(hjust = 0.5))
g <- g + labs(y = "Size of property and crop loss", x = "Event Type")
g <- g + coord_flip()
print(g)
```


Total economic loss in USA by weather events in 1996−2011
FLOOD
HURRICANE/TYPHOON
STORM SURGE
TORNADO
HAIL
FLASH FLOOD
HURRICANE
DROUGHT
TROPICAL STORM
HIGH WIND
Event Type
0.0e+00
5.0e+10
1.0e+11
1.5e+11
Size of property and crop loss