Effects of severe weather events in US on health and economy

Synopsis

In this analysis I explore the effects of severe weather events in the US on health and economy.

Data Processing

First I load the StormData.

```
#Download and unzip the data file
#https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2
address <- "https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2"
address <- sub("^https", "http", address) #https does not work
zipname <- "StormData.csv.bz2"
download.file(address,zipname)

#No need to unzip, R can is smart enough to read a zipped csv
stormdata <- read.csv(zipname)

#housekeeping - remove the zip as it is no longer needed
file.remove(zipname)
```

```
## [1] TRUE
```

```
#housekeeping
rm(address, zipname)
```

Some more data preparation. I add a column called HEALTH that contains the number of people affected. I include both the fatalities and the injuries, i.e. the fataly and non-fataly injured. Secondly I add a column for the damages in dollars. For this I need to combine the columns with numbers and columns with alphabetical indictation, for the severity, “K” for thousands of dollars, “M” for millions, “B” for billions. Thirdly, I add a column for year. The first years a scarcely filled with data and inculde only data for a small number of event types. From the nineties onwards, there is a more complete set.

```
#Add a health column
#This includes all those that are affected, i.e. killed or injured
stormdata$HEALTH <- stormdata$FATALITIES + stormdata$INJURIES

#Add a column that adds up the prop and crop damage
#Start with all zeros and then add propdmg and cropdmg multiplied by
#the correct number depending on "K", "M" or "B"
K <- 1000
M <- 1000000
B <- 1000000000

#Set damage to all zeros
stormdata$DAMAGE <- 0
```

```
#Add PROPDMG thousands, millions and billions
stormdata$DAMAGE[toupper(stormdata$PROPDMGEXP)=="K"] <-
  stormdata$DAMAGE[toupper(stormdata$PROPDMGEXP)=="K"] +
  stormdata$PROPDMG[toupper(stormdata$PROPDMGEXP)=="K"] * K
stormdata$DAMAGE[toupper(stormdata$PROPDMGEXP)=="M"] <-
  stormdata$DAMAGE[toupper(stormdata$PROPDMGEXP)=="M"] +
  stormdata$PROPDMG[toupper(stormdata$PROPDMGEXP)=="M"] * M
stormdata$DAMAGE[toupper(stormdata$PROPDMGEXP)=="B"] <-
  stormdata$DAMAGE[toupper(stormdata$PROPDMGEXP)=="B"] +
  stormdata$PROPDMG[toupper(stormdata$PROPDMGEXP)=="B"] * B

#Add CROPDMG thousands, millions and billions
stormdata$DAMAGE[toupper(stormdata$CROPDMGEXP)=="K"] <-
  stormdata$DAMAGE[toupper(stormdata$CROPDMGEXP)=="K"] +
  stormdata$CROPDMG[toupper(stormdata$CROPDMGEXP)=="K"] * K
stormdata$DAMAGE[toupper(stormdata$CROPDMGEXP)=="M"] <-
  stormdata$DAMAGE[toupper(stormdata$CROPDMGEXP)=="M"] +
  stormdata$CROPDMG[toupper(stormdata$CROPDMGEXP)=="M"] * M
stormdata$DAMAGE[toupper(stormdata$CROPDMGEXP)=="B"] <-
  stormdata$DAMAGE[toupper(stormdata$CROPDMGEXP)=="B"] +
  stormdata$CROPDMG[toupper(stormdata$CROPDMGEXP)=="B"] * B

#house keeping
rm(K, M , B)

#Use library lubridate
#Only the data for more recent years is relevant
library(lubridate)

#Add a YEAR colum to the data
stormdata$YEAR <- year(mdy_hms(stormdata$BGN_DATE))
#Take into account only the years for which we have more complete data
#I set the year to be larger or equal to 1994
complete_year <- 1994
stormdata <- stormdata[which(stormdata$YEAR >= complete_year),]

#house keeping
rm(complete_year)
```

Explore the health data

```
#aggregate over health effects i.e. injuries
#fatalities do no have an effect on population health
#use !is.na or na.rm to exclude NA's
health_count <- aggregate(HEALTH ~ EVTYPE, data=stormdata, function(x) length( !is.na(x)))
health_sum <- aggregate(HEALTH ~ EVTYPE, data=stormdata, function(x) sum(x, na.rm = TRUE))
health_mean <- aggregate(HEALTH ~ EVTYPE, data=stormdata, function(x) mean(x, na.rm = TRUE))

names(health_count)[2] <- "number_events"
names(health_sum)[2] <- "number_affected"
names(health_mean)[2] <- "mean_number_affected"
```

```
#combine in one dataframe
health <- data.frame(health_count,
                     health_sum[2],
                     health_mean[2]
                     )

#house keeping
rm(health_count, health_sum, health_mean)

#calculate top10
health_top10 <- health[order(health$number_affected, decreasing = TRUE),][1:10,]

health_top10
```

```
##                EVTYPE number_events number_affected mean_number_affected
## 781           TORNADO         25274           24164           0.95608135
## 115    EXCESSIVE HEAT          1678            8428           5.02264601
## 154             FLOOD         24906            7228           0.29021119
## 425         LIGHTNING         15287            5910           0.38660300
## 804         TSTM WIND        128970            3872           0.03002249
## 254              HEAT           759            3025           3.98550725
## 139       FLASH FLOOD         53396            2705           0.05065923
## 398         ICE STORM          1972            2057           1.04310345
## 712 THUNDERSTORM WIND         82482            1609           0.01950729
## 918      WINTER STORM         11403            1493           0.13093046
```

Explore the economic damage data

```
#aggregate over economic damage
#fatalities do no have an effect on population health
#use !is.na or na.rm to exclude NA's
damage_count <- aggregate(DAMAGE ~ EVTYPE, data=stormdata, function(x) length( !is.na(x)))
damage_sum <- aggregate(DAMAGE ~ EVTYPE, data=stormdata, function(x) sum(x, na.rm = TRUE))
damage_mean <- aggregate(DAMAGE ~ EVTYPE, data=stormdata, function(x) mean(x, na.rm = TRUE))

names(damage_count)[2] <- "number"
names(damage_sum)[2] <- "damage"
names(damage_mean)[2] <- "mean_damage"

#combine in one dataframe
damage <- data.frame(damage_count,
                     damage_sum[2],
                     damage_mean[2]
                     )

#house keeping
rm(damage_count,damage_sum, damage_mean)

#calculate top10
damage_top10 <- damage[order(damage$damage, decreasing = TRUE),][1:10,]
```

```
damage_top10
```

```
##                EVTYPE number       damage  mean_damage
## 154             FLOOD  24906 149686551250   6010059.88
## 381 HURRICANE/TYPHOON     88  71913712800 817201281.82
## 620       STORM SURGE    253  43193541000 170725458.50
## 781           TORNADO  25274  25982192520   1028020.60
## 224              HAIL 222616  18317742770     82284.04
## 139       FLASH FLOOD  53396  17119054960    320605.57
## 83            DROUGHT   2478  14968172000   6040424.54
## 373         HURRICANE    171  14604229010  85404848.01
## 398         ICE STORM   1972   8854491310   4490107.16
## 794    TROPICAL STORM    686   8381226550  12217531.41
```

Results