Severe Weather Events and his impact on USA

Carlos Pinilla

Wednesday, June 17, 2015

Synopsis

This article has the aim to evidence, on the basis of an official publication of the National Oceanic Atmospheric Administration (NOAA), called "Storm Data" published on August 17, 2007, which are types of natural disasters that has suffered the United States and has had the worst consequences respect to population health and economics. The results shows that the principal event impact the population health and economy are the tornados and the principal event impact the economy in the crops is the heat. To guide the lector to the description of the analisys performed, below will find the logics steps followed to achieve the results indicated.

Data Procesising

Libraries ang general setup

```
echo = FALSE  # Avoid the warning message of library and versions
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.1.3

library(plyr)

## Warning: package 'plyr' was built under R version 3.1.3
```

Download the data from the NOAA

The data can be downloaded from the following link Storm Data

```
#wd <-
setwd("D:/cpinilla/personal/DataScientific/Reproducible_Research/Assessme
nt2")
if ( !"stormData.csv.bz2" %in% dir()){
        print("Let me download the file")

download.file("http://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStor
mData.csv.bz2", destfile = "stormData.csv.bz2")}
if (!"StormData" %in% ls()){
        print("Loading the data")
        StormData <- read.csv("stormData.csv.bz2")}

## [1] "Loading the data"
```

Data procesing

Analizing all the event type, there are many cases disgregated by the spelling (i.e Cold,cold,Cold) or for likely events (i.e Cold, frezzeng, hypotermia) So first it will be agregated a column with a key word that group all the cases in the likely and misspelling events

```
StormData[grepl("precipitation|rain|hail|drizzle|wet|percip|burst|depress
ion|fog|wall cloud",StormData$EVTYPE, ignore.case = TRUE), "GROUP.EVENT"]
<- "PRECIPITATION"
StormData[grepl("wind|wnd|hurricane|typhoon|torn|spout|funnel",StormData$
EVTYPE, ignore.case = TRUE), "GROUP.EVENT"] <- "TORNADO"
StormData[grepl("slide|erosion|slump|aval",StormData$EVTYPE, ignore.case
= TRUE), "GROUP.EVENT"] <- "EROSION"
StormData[grepl("warmth|warm|heat|dry|hot|drought|hyperthermia|temperatur
e record|record temperature|record high",StormData$EVTYPE, ignore.case =
TRUE), "GROUP.EVENT"] <- "HEAT"
StormData[grepl("cold|cool|ice|icy|frost|freeze|snow|winter|wintry|winter
y|blizzard|chill|freezing|glaze|sleet|hypo",StormData$EVTYPE, ignore.case
= TRUE), "GROUP.EVENT"] <- "COLD"
StormData[grepl("floo|surf|blow-out|swells|fld|dam
break",StormData$EVTYPE, ignore.case = TRUE), "GROUP.EVENT"] <- "FLOOD"
StormData[grepl("seas |high
water|tide|tsunami|wave|current|marine|drowning",StormData$EVTYPE,
ignore.case = TRUE), "GROUP.EVENT"] <- "SEAS"
StormData[grepl("dust|saharan",StormData$EVTYPE, ignore.case = TRUE),
"GROUP.EVENT"] <- "DUST"
StormData[grepl("tstm|thunderstorm|lightning",StormData$EVTYPE,
ignore.case = TRUE), "GROUP.EVENT"] <- "THUNDER"
StormData[grepl("fire|smoke|volcanic",StormData$EVTYPE, ignore.case =
TRUE), "GROUP.EVENT"] <- "FIRE"
```

The two main question this article wants to address are:

1. Across the United States, which types of events (as indicated in the EVTYPE variable) are most harmful with respect to population health?

2. Across the United States, which types of events have the greatest economic consequences?

So after revise the data structure of the NOAA report, it were selected 9 variables from 38. The variables are:

+BGN_DATE : Date when was register the ocurrence of the event. +EVTYPE : Which type of event occours +GROUP.EVENT: Group all similar events +FATALITIES: Number of deseases +INJURIES : Number of injuries +PROPDMG : Property damage estimates +PROPDMGEXP: Alphabetical characters used to signify magnitude (i.e K = 1000) +CROPDMG : Crop damage estimates +CROPDMGEXP: Alphabetical characters used to signify magnitude (i.e K = 1000)

```
Variables.used <- c("BGN_DATE", "EVTYPE","GROUP.EVENT", "FATALITIES",
"INJURIES", "PROPDMG", "PROPDMGEXP", "CROPDMG", "CROPDMGEXP")
  StormData <- StormData[, Variables.used]
```

Now it will be obteined the events which has reported the maximun number of fatalities and Injuries

```
Fatal.Table <- aggregate(FATALITIES ~ GROUP.EVENT, data = StormData, FUN
= sum)
Fatal.10 <- head(arrange(Fatal.Table, desc(FATALITIES)), n = 10)

Injur.Table <- aggregate(INJURIES ~ GROUP.EVENT, data = StormData, FUN =
sum)
Injur.10 <- head(arrange(Injur.Table, desc(INJURIES)), n = 10)
```

For summarize de amount in US dollars, we have to translate the PROPDMGEXP and CROPDMGEXP and multiply the units in the propertie damage and the exponential plus the multiplication of the unit in the crop damage for the exponential and register the result in one column.

```
StormData$PROPDMGEXP <- as.character(StormData$PROPDMGEXP)
StormData$CROPDMGEXP <- as.character(StormData$CROPDMGEXP)
# A function to convert K, M, B units in --EXP columns
Translate <- function(x){
  x[x==""]  <- "1"
  x[x=="?"] <- "1"
  x[x=="K"] <- "1000"
  x[x=="k"] <- "1000"
  x[x=="M"] <- "1000000"
  x[x=="m"] <- "1000000"
  x[x=="B"] <- "1000000000"
  return(as.numeric(x))
}

StormData$PROPDMGEXP <- Translate(StormData$PROPDMGEXP)

## Warning in Translate(StormData$PROPDMGEXP): NAs introducidos por
coerción

StormData$CROPDMGEXP <- Translate(StormData$CROPDMGEXP)

StormData <- mutate(StormData,
Sum.Damage=PROPDMG*PROPDMGEXP+CROPDMG*CROPDMGEXP)
StormData <- mutate(StormData, Prop.Damage=PROPDMG*PROPDMGEXP)
StormData <- mutate(StormData, Crop.Damage=CROPDMG*CROPDMGEXP)
```

Now it must be summarized the total damage by each event

```
Prop.Table <- aggregate(Prop.Damage ~GROUP.EVENT, data=StormData, FUN =
sum)
Prop.10 <- head(arrange(Prop.Table,desc(Prop.Damage)),n=10)
```

```
Crop.Table <- aggregate(Crop.Damage ~GROUP.EVENT, data=StormData, FUN =
sum)
Crop.10 <- head(arrange(Crop.Table,desc(Crop.Damage)),n=10)
Damage.Table <- aggregate(Sum.Damage ~ GROUP.EVENT , data=StormData,
FUN=sum)
Damage.10 <- head(arrange(Damage.Table,desc(Sum.Damage)),n=10)
```

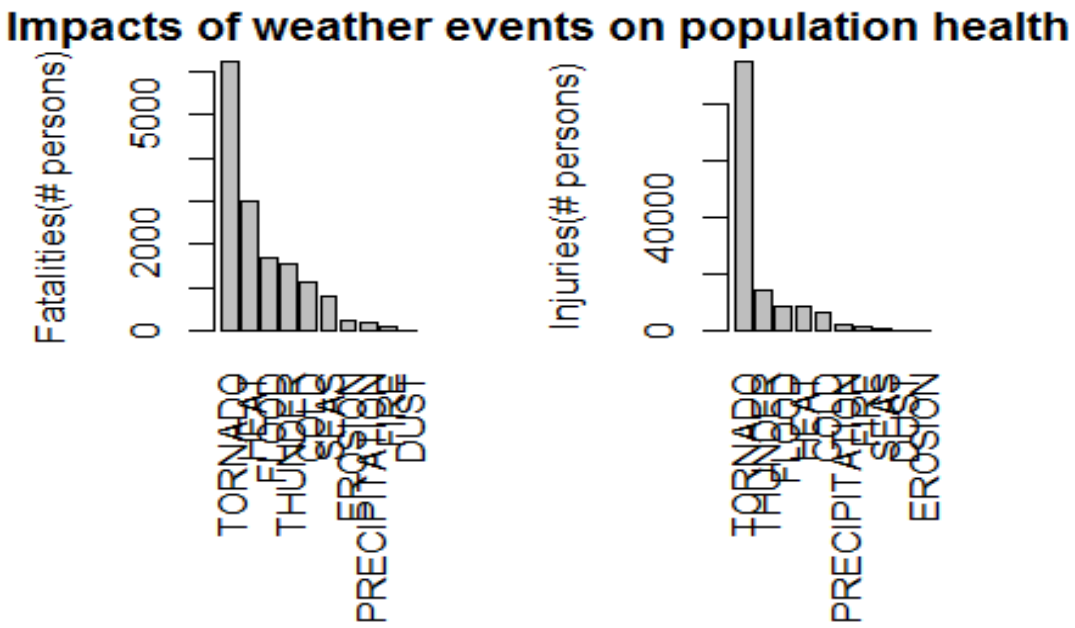
Results

Impact on peoples health

From the data, we realize that there are 10 categories where it can be classified almost all the natural events recorded in the data from the NOAA.

The natural catastrofies that has affect the population health organized from harmfully to lesser ones are

```
par(mfrow = c(1, 2), mar = c(10, 4, 0.5, 2.5), oma = c(1.5, 2, 1, 1))
barplot(Fatal.10$FATALITIES, names.arg = Fatal.10$GROUP.EVENT, ylab =
"Fatalities(# persons)", las = 3)
barplot(Injur.10$INJURIES, names.arg = Injur.10$GROUP.EVENT, ylab =
"Injuries(# persons)", las = 3)
title(main = "Impacts of weather events on population health", outer =
TRUE)
```

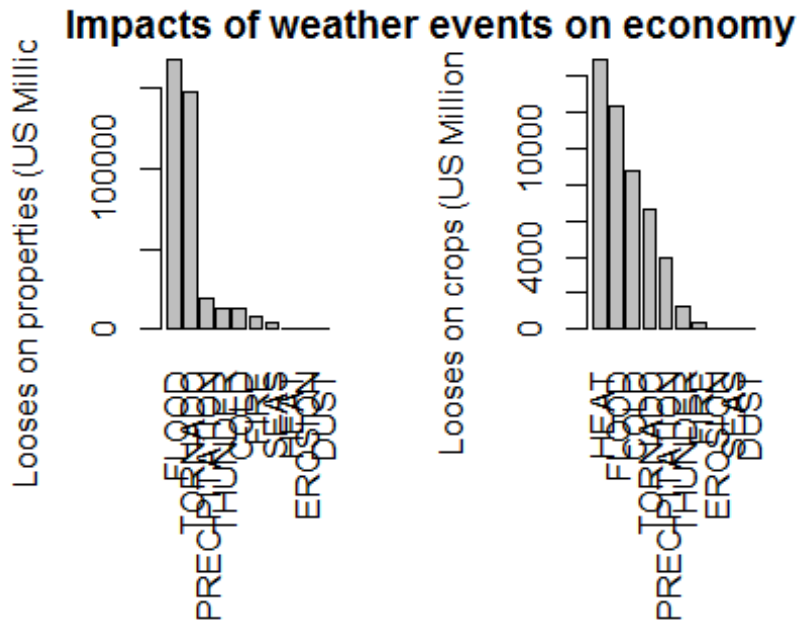


####. #### Fig 1 Impacts of natural events on population health ####.

Impact on economy

In the same way as the previus analisys, here are presented the natural events that has affect the USA economy from 1950 to 2011 organized from the expensive ones.

```
par(mfrow = c(1, 2), mar = c(10, 4, 0.5, 2.5), oma = c(1.5, 2, 1, 1))
barplot(Prop.10$Prop.Damage/1000000, names.arg = Prop.10$GROUP.EVENT,
ylab = "Looses on properties (US Million dollars)", las = 3)
barplot(Crop.10$Crop.Damage/1000000, names.arg = Crop.10$GROUP.EVENT,
ylab = "Looses on crops (US Million dollars)", las = 3)
title(main = "Impacts of weather events on economy", outer = TRUE)
```



####. ####

Fig 1 Impacts of natural events on economy ####.