Impact of Severe Weather Events on Health and Economic Factors in U.S. for 1950-2011

Tjeerd Luykx

Wednesday, July 22, 2015

1. Synopsis:

Storms and other severe weather events are able to 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 report 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 f.atalities, injuries, and property damage.

The basic goal of this report is to explore the NOAA Storm Database and answer some basic questions about severe weather events. The analysis can consists of tables, figures, or other summaries. The following questions are addressed in this report:

1. Across the United States, which types of events are most harmful with respect to population health?
2. Across the United States, which types of events have the greatest economic consequences?

2. Data Processing:

2.1 Package Load:

The following packages are used to conduct loading, computation and graphical reporting. For displaying purposes also digit and width options are set below.

```
library(knitr)
library(RCurl)
library(R.utils)
library(dplyr)
library(ggplot2)
library(lubridate)


options(digits = 7)
opts_chunk$set(fig.width=10)
```

2.2 Weather Event Data Load:

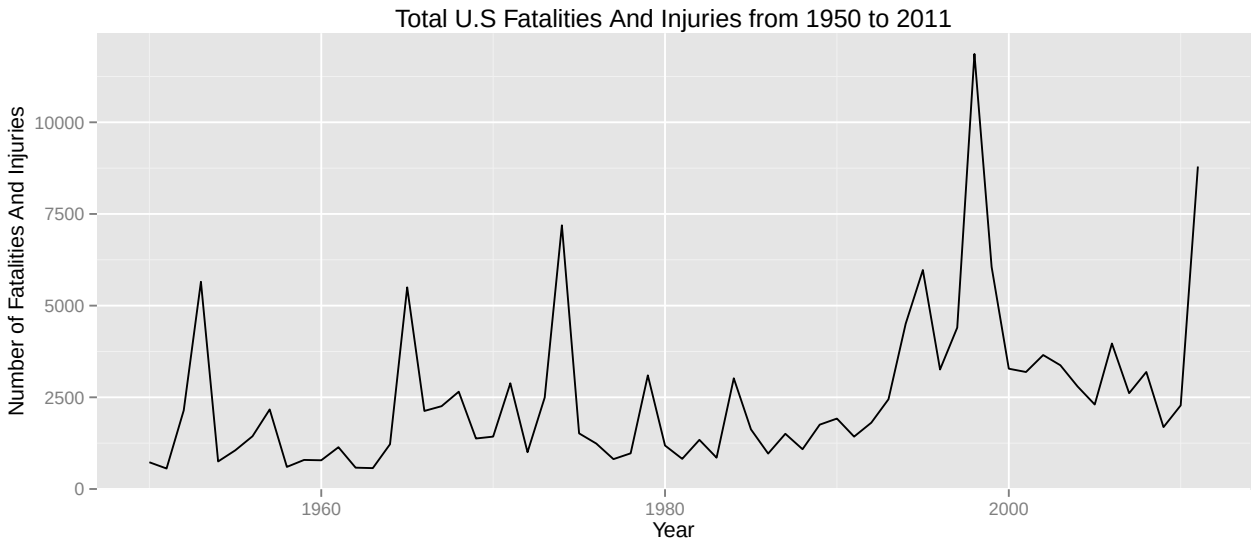
The data processing in this report consists of downloading the dataset from the URL, unzipping the file and creating an adequate data frame for analysis. Prior to the dataset download a special directory is created named data.

```
# Creating data directory:
if (!file.exists("./data")){
      dir.create("./data")
```

```
      }

# Downloading dataset:
  if(!file.exists("./data/StormData.csv.bz2")){
    FileURL <- "https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2"
    FileDir <- "./data/StormData.csv.bz2"
    FileBin <- getBinaryURL(FileURL, ssl.verifypeer = 0L, followlocation = 1L)
    FileDest <- file(FileDir, open = "wb")
    writeBin(FileBin, FileDest)
    close(FileDest)
    }

# Unzipping bz2 file and reformat to csv:
  if(!file.exists("./data/StormData.csv")){
    FilePath <- "./data/StormData.csv.bz2"
    FileDir <- "./data/StormData.csv"
    bunzip2(FilePath,FileDir, overwrite = TRUE, remove = FALSE)
  }

# Create data frame from weather data
W.Data <- read.csv("./data/StormData.csv")
```

2.3 Filtering Dataset:

```
W.Data.s <- select(W.Data, BGN_DATE, EVTYPE, FATALITIES, INJURIES, PROPDMG, PROPDMGEXP, CROPDMG, CROPDM
W.Data.s$Year <- as.numeric(format(as.Date(W.Data.s$BGN_DATE, format="%m/%d/%Y %H:%M:%S"),"%Y"))

W.Data.agg <- aggregate(formula = cbind(FATALITIES + INJURIES) ~ Year, data = W.Data.s, FUN = sum)

g1 <- ggplot(W.Data.agg , aes(Year, V1))
g1 <- g1 + geom_line() +
        xlab("Year") +
        ylab(expression("Number of Fatalities And Injuries")) +
        ggtitle("Total U.S Fatalities And Injuries from 1950 to 2011")
print(g1)
```




```
W.Data.f <- filter(W.Data.s, Year >= "1990")
```

2.4 Formatting Dataset:

```
ExpForm <- function(exp) {
    if (exp %in% c('h', 'H'))
        return(2)
    else if (exp %in% c('k', 'K'))
        return(3)
    else if (exp %in% c('m', 'M'))
        return(6)
    else if (exp %in% c('b', 'B'))
        return(9)
    else if (!is.na(as.numeric(exp)))
        return(as.numeric(exp))
    else if (exp %in% c('', '-', '?', '+'))
        return(0)
    else {
        stop("Invalid Value.")
    }
}

PropDmgExp <- sapply(W.Data.f$PROPDMGEXP, FUN=ExpForm)
W.Data.f$PropDmg <- W.Data.f$PROPDMG * (10 ** PropDmgExp)
CropDmgExp <- sapply(W.Data.f$CROPDMGEXP, FUN=ExpForm)
W.Data.f$CropDmg <- W.Data.f$CROPDMG * (10 ** CropDmgExp)
```

2.4 Inflation Correction:

```
CPI.m <- read.csv(file="./data/CPIAUCSL.csv", header = TRUE)
CPI.m $cpi_year <- year(CPI.m$DATE)
CPI.y <- CPI.m %>% group_by(cpi_year) %>% summarize(cpi = mean(VALUE))
```

```
CPI.y <- as.data.frame(CPI.y)
CPI.y$Adj <- CPI.y$cpi[CPI.y$cpi_year == 2013]/CPI.y$cpi
colnames(CPI.y) <- c("Year","CPI","CPIConv")

W.Data.m <- full_join(W.Data.f, CPI.y, by='Year')
W.Data.m$TotEcoDmg <- (W.Data.m$PropDmg  + W.Data.m$CropDmg) * W.Data.m$CPIConv
```

2.4 Reformatting Event Types:

```
W.Data.m$EVTYPE <- gsub(("HURRICANE/TYPHOON|HURRICANE OPAL|HURRICANE OPAL/HIGH WINDS|HURRICANE EMILY|TY

W.Data.m$EVTYPE <- gsub(("TSTM WIND|TSTM WIND|SEVERE THUNDERSTORM WINDS|THUNDERSTORM WINDS"),"THUNDERST

W.Data.m$EVTYPE <- gsub(("HEAVY RAIN/SEVERE WEATHER|EXCESSIVE RAINFALL|UNSEASONAL RAIN|HEAVY RAINS"),"H

W.Data.m$EVTYPE <- gsub(("STORM SURGE/TIDE"),"STORM SURGE", ignore.case = TRUE, x = W.Data.m$EVTYPE)

W.Data.m$EVTYPE <- gsub(("WILD/FOREST FIRE|WILDFIRES|WILD FIRES"),"WILDFIRE", ignore.case = TRUE, x = W

W.Data.m$EVTYPE <- gsub(("EXCESSIVE HEAT|HEAT WAVE|EXTREME HEAT|UNSEASONABLY WARM|RECORD/EXCESSIVE HEAT

W.Data.m$EVTYPE <- gsub(("EXTREME COLD|FROST/FREEZE|FROST|Early Frost|DAMAGING FREEZE|RECORD COLD|COLD/

W.Data.m$EVTYPE <- gsub(("HIGH WINDS|HIGH WIND|BLOWING WIND|STRONG WINDS|STRONG WIND"),"WIND", ignore.c

W.Data.m$EVTYPE <- gsub(("FLASH FLOODING|FLASH FLOOD/FLOOD|FLOOD/FLASH FLOOD"),"FLASH FLOOD", ignore.cas

W.Data.m$EVTYPE <- gsub(("SMALL HAIL"),"HAIL", ignore.case = TRUE, x = W.Data.m$EVTYPE)

W.Data.m$EVTYPE <- gsub(("RIVER FLOODING"),"RIVER FLOOD", ignore.case = TRUE, x = W.Data.m$EVTYPE)

W.Data.m$EVTYPE <- gsub(("FLOODING|MAJOR FLOOD"),"FLOOD", ignore.case = TRUE, x = W.Data.m$EVTYPE)

W.Data.m$EVTYPE <- gsub(("COASTAL FLOODING|COASTAL FLOODING/EROSION|COASTAL FLOODING/EROSION|Erosion/Cs

W.Data.m$EVTYPE <- gsub(("TROPICAL STORM GORDON|TROPICAL STORM JERRY"),"TROPICAL STORM", ignore.case = 

W.Data.m$EVTYPE <- gsub(("DENSE FOG"),"FOG", ignore.case = TRUE, x = W.Data.m$EVTYPE)

W.Data.m$EVTYPE <- gsub(("RIP CURRENTS"),"RIP CURRENT", ignore.case = TRUE, x = W.Data.m$EVTYPE)

W.Data.m$EVTYPE <- gsub(("HEAVY SURF|HEAVY SURF/HIGH SURF"),'HIGH SURF', ignore.case = TRUE, x = W.Data

W.Data.m$EVTYPE <- gsub(("WATERSPOUT/TORNADO"),"WATERSPOUT", ignore.case = TRUE, x = W.Data.m$EVTYPE)

W.Data.m$EVTYPE <- gsub(("WINTRY MIX|WINTER WEATHER MIX|WINTER WEATHER/MIX"),"WINTER WEATHER", ignore.c

W.Data.m$EVTYPE <- gsub(("WINTER STORMS"),"WINTER STORM", ignore.case = TRUE, x = W.Data.m$EVTYPE)

W.Data.m$EVTYPE <- gsub(("MARINE TSTM WIND"),"MARINE THUNDERSTORM WIND", ignore.case = TRUE, x = W.Data
```

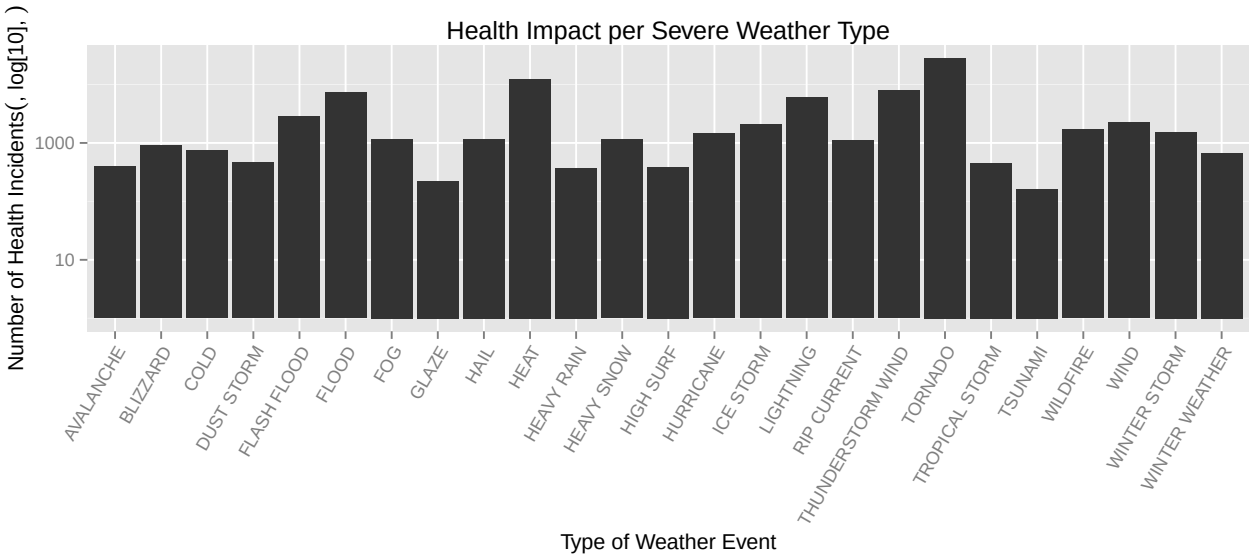
2.5 Summation Injuries and Fatalities:

```
W.Data.m$TotHeaDmg <- W.Data.m$INJURIES + W.Data.m$FATALITIES
```

3. Results:

3.1. Healt Impact:

```
HeaImp <- head(arrange(aggregate(TotHeaDmg ~ EVTYPE,W.Data.m,sum,na.rm=TRUE),desc(TotHeaDmg)),n=25)
colnames(HeaImp) <- c("EventType","HealthIncidents")

p1 <- ggplot(data=HeaImp, aes(x=EventType, y=HealthIncidents))
p1 <- p1 + geom_bar(stat="identity") +
        labs(x="", y="Number of Health Incidents")
p2 <- p1 + scale_y_log10() +
        theme(axis.text.x = element_text(angle = 60, hjust = 1)) +
        labs(y=expression(paste("Number of Health Incidents",(", log[10], ")))) +
        xlab("Type of Weather Event") +
        ggtitle("Health Impact per Severe Weather Type")
p1 <- p1 + theme(axis.text.x = element_blank(), axis.ticks.x = element_blank())
print(p2)
```



3.1. Economic Impact:

```
AggEcoDmg <- head(arrange(aggregate(TotEcoDmg ~ EVTYPE,W.Data.m,sum,na.rm=TRUE),desc(TotEcoDmg)),n=25)
colnames(AggEcoDmg) <- c("EventType","TotEcoDmg")

p3 <- ggplot(data=AggEcoDmg, aes(x=EventType, y=TotEcoDmg))
p3 <- p3 + geom_bar(stat="identity") +
        labs(x="", y="Property and Crop Damage in U.S. Dollar")
```

```
p4 <- p3 + scale_y_log10() +
        theme(axis.text.x = element_text(angle = 60, hjust = 1)) +
        labs(y=expression(paste("Property and Crop Damage in U.S. Dollar",(", log[10], ")))) +
        xlab("Type of Weather Event") +
        ggtitle("Economic Impact per Severe Weather Type")
p3 <- p3 + theme(axis.text.x = element_blank(), axis.ticks.x = element_blank())
print(p4)
```