Catastrophe

Author: Aditya Gudal

Impact Of Natural Catastrophes On The Economy And The Public Health Of The United States

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.

Abstract

This report consists in analyzing the NOAA storm database containing data on extreme climate events. This data was collected during the period from 1950 through 2011. The purpose of this analysis is to answer the following two questions:

Across the United States, which types of events (as indicated in the EVTYPE variable) are most harmful with respect to population health? Across the United States, which types of events have the greatest economic consequences? I used the weather events specified in the documentation (paragraphs 7.1 - 7.48). Main conclusions of the study:

Tornado is the most hazordous climate event with more than 5600 deaths and 91400 injuries. Floods have caused the most significant economic damage - more than 157 billion USD.

Basic libraries needed.

```
options(warn=-1)
options(scipen = 1)   # Turn off scientific notations for numbers
library(grid)
library(ggplot2)
library(plyr)
require(gridExtra)
```

```
## Loading required package: gridExtra
```

Data processing

We read the csv file.

```
library(readr)
storm <- read_csv("C:/Users/Lenovo/Desktop/New folder
(2)/repdata%2Fdata%2FStormData.csv")

## Parsed with column specification:
## cols(
##   .default = col_character(),
##   STATE__ = col_double(),
##   COUNTY = col_double(),
##   BGN_RANGE = col_double(),
##   COUNTY_END = col_double(),
##   END_RANGE = col_double(),
##   LENGTH = col_double(),
##   WIDTH = col_double(),
##   F = col_integer(),
##   MAG = col_double(),
##   FATALITIES = col_double(),
##   INJURIES = col_double(),
##   PROPDMG = col_double(),
##   CROPDMG = col_double(),
##   LATITUDE = col_double(),
##   LONGITUDE = col_double(),
##   LATITUDE_E = col_double(),
##   LONGITUDE_ = col_double(),
##   REFNUM = col_double()
## )

## See spec(...) for full column specifications.

dim(storm)

## [1] 902297     37
```

Extract data corresponding to the 48 events as described in the documentation paragraphs 7.1 through 7.48

1.Vector of 48 events as defined in the documentation:

```
events <- c("Astronomical Low Tide", "Avalanche", "Blizzard", "Coastal
Flood", "Cold/Wind Chill", "Debris Flow", "Dense Fog", "Dense Smoke",
"Drought", "Dust Devil", "Dust Storm", "Excessive Heat", "Extreme cold/Wind
Chill", "Flash Flood", "Flood", "Freezing", "Frost/Freeze", "Funnel Cloud",
"Hail", "Heat", "Heavy Rain", "Heavy Snow", "High Surf", "High Wind",
"Hurricane/Typhoon", "Ice Storm", "Lakeshore Flood", "Lake-Effect Snow",
"Lightning", "Marine Hail", "Marine High Wind", "Marine Strong Wind", "Marine
Thunderstorm Wind", "Rip Current", "Seiche", "Sleet", "Storm Tide", "Strong
Wind", "Thunderstorm Wind", "Tornado", "Tropical Depression", "Tropical
Storm", "Tsunami", "Volcanic Ash", "Waterspout", "Wildfire", "Winter Storm",
"Winter Weather")
```

2.Some events are combined events separated with a slash (e.g 'Hurricane/Typhoon'). I will use regular expressions to extract either a combined event (Hurricane/Typhoon) or any part of it (Hurricane or Typhoon).

```
events_regex <- c("Astronomical Low Tide|Low Tide", "Avalanche", "Blizzard",
"Coastal Flood", "Cold/Wind Chill", "Debris Flow", "Dense Fog", "Dense
Smoke", "Drought", "Dust Devil", "Dust Storm", "Excessive Heat", "Extreme
cold/Wind Chill|Extreme Cold|Wind Chill", "Flash Flood", "Flood", "Freezing",
"Frost/Freeze|Frost|Freeze", "Funnel Cloud", "Hail", "Heat", "Heavy Rain",
"Heavy Snow", "High Surf", "High Wind",
"Hurricane/Typhoon|Hurricane|Typhoon", "Ice Storm", "Lakeshore Flood", "Lake-
Effect Snow", "Lightning", "Marine Hail", "Marine High Wind", "Marine Strong
Wind", "Marine Thunderstorm Wind|Marine tstm Wind", "Rip Current", "Seiche",
"Sleet", "Storm Tide", "Strong Wind", "Thunderstorm Wind|tstm wind",
"Tornado", "Tropical Depression", "Tropical Storm", "Tsunami", "Volcanic
Ash", "Waterspout", "Wildfire", "Winter Storm", "Winter Weather")
```

3.The next step is to extract rows corresponding to the event from the documentation, I will also choose the columns which are relevant to our analysis: *EVTYPE -> Type of event* FATALITIES -> Number of fatalities *INJURIES -> Number of injuries* PROPDMG -> Amount of property damage in orders of magnitude *PROPDMGEXP -> Order of magnitude for property damage (e.g. K for thousands)* CROPDMG -> Amount of crop damage in orders of magnitude *PROPDMGEXP -> Order of magnitude for crop damage (e.g. M for millions)

```
options(scipen = 999)  # force fixed notation of numbers instead of
scientific
cleandata <- data.frame(EVTYPE = character(0), FATALITIES = numeric(0),
INJURIES=numeric(0), PROPDMG = numeric(0), PROPDMGEXP = character(0), CROPDMG
= numeric(0), CROPDMGEXP = character(0))
for (i in 1:length(events)) {
    rows <- storm[grep(events_regex[i], ignore.case = TRUE, storm$EVTYPE), ]
    rows <- rows[, c("EVTYPE", "FATALITIES", "INJURIES", "PROPDMG",
"PROPDMGEXP","CROPDMG",     "CROPDMGEXP")]
    CLEANNAME <- c(rep(events[i], nrow(rows)))
    rows <- cbind(rows, CLEANNAME)
    cleandata <- rbind(cleandata, rows)
}
```

4.Take into account the order of magnitude of property and crop damage (H = hundreds, K = thousands, M = millions, B= billions)

5.Compute combined economic damage (property damage + crops damage)

```
# multiply property and crops damage by 10 raised to the power of the
exponent
suppressWarnings(cleandata$PROPDMG <- cleandata$PROPDMG *
10^as.numeric(cleandata$PROPDMGEXP))
suppressWarnings(cleandata$CROPDMG <- cleandata$CROPDMG *
10^as.numeric(cleandata$CROPDMGEXP))
# compute combined economic damage (property damage + crops damage)
```

```
suppressWarnings(TOTECODMG <- cleandata$PROPDMG + cleandata$CROPDMG)
cleandata <- cbind(cleandata, TOTECODMG)
# delete 'PROPDMGEXP' and 'CROPDMGEXP'columns which have become unnecessary
after conversion
cleandata <- cleandata[, c("EVTYPE", "FATALITIES", "INJURIES", "PROPDMG",
"CROPDMG", "CLEANNAME", "TOTECODMG")]
```

Observations

Question 01 : Across the United States, which types of events are most harmful with respect to population health?

Fatalities and injuries *As for the impact on public health, we have got two sorted lists of severe weather events below by the number of people badly affected. Aggregate data for fatalities.

```
fatalities <- aggregate(FATALITIES ~ CLEANNAME, data = cleandata, FUN = sum)
fatalities <- fatalities[order(fatalities$FATALITIES, decreasing = TRUE), ]
# 10 most harmful causes of fatalities
MaxFatalities <- fatalities[1:10, ]
print(MaxFatalities)
```

```
##                  CLEANNAME FATALITIES
## 38                 Tornado       5661
## 19                    Heat       3138
## 11          Excessive Heat       1922
## 14                   Flood       1525
## 13             Flash Flood       1035
## 28               Lightning        817
## 37       Thunderstorm Wind        753
## 33             Rip Current        577
## 12 Extreme cold/Wind Chill        382
## 23               High Wind        299
```

Aggregate data for injuries.

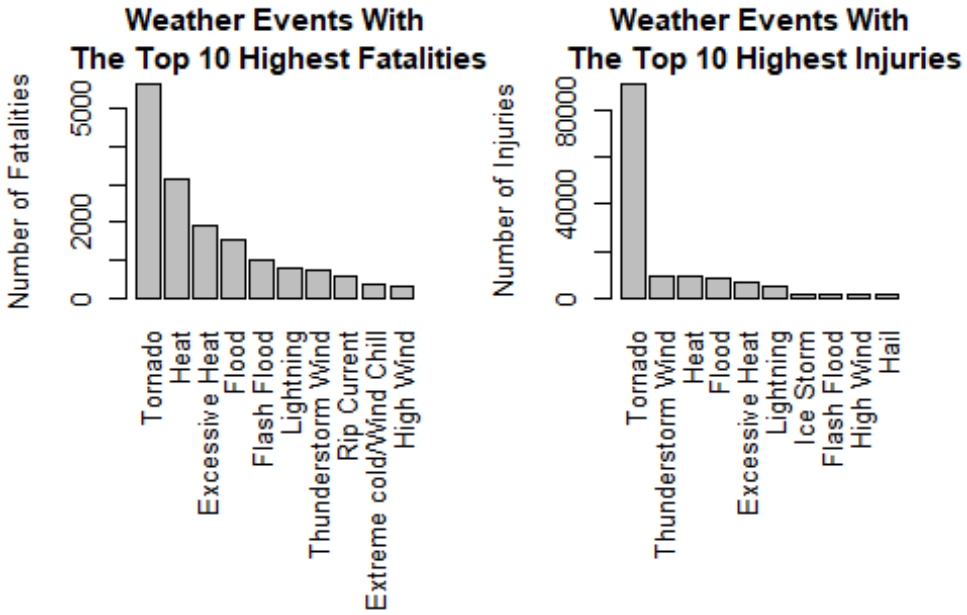
```
injuries <- aggregate(INJURIES ~ CLEANNAME, data = cleandata, FUN = sum)
injuries <- injuries[order(injuries$INJURIES, decreasing = TRUE), ]
# 10 most harmful causes of injuries
MaxInjuries <- injuries[1:10, ]
print(MaxInjuries)
```

```
##            CLEANNAME INJURIES
## 38           Tornado    91407
## 37 Thunderstorm Wind     9493
## 19              Heat     9224
## 14             Flood     8604
## 11    Excessive Heat     6525
## 28         Lightning     5232
```

```
## 25         Ice Storm     1992
## 13       Flash Flood     1802
## 23         High Wind     1523
## 18              Hail     1467
```

The graphs showing total fatalities and injuries due to weather conditions.

```
par(mfrow = c(1, 2), mar = c(15, 4, 3, 2), mgp = c(3, 1, 0), cex = 0.8)
barplot(MaxFatalities$FATALITIES, las = 3, names.arg =
MaxFatalities$CLEANNAME, main = "Weather Events With\n The Top 10 Highest
Fatalities", ylab = "Number of Fatalities", col = "grey")
barplot(MaxInjuries$INJURIES, las = 3, names.arg = MaxInjuries$CLEANNAME,
main = "Weather Events With\n The Top 10 Highest Injuries", ylab = "Number of
Injuries", col = "grey")
```



Based on the above histograms, we find that Tornado and Heat had caused most fatalities and Tornado had caused most injuries in the United States from 1995 to 2011.

Note: I decided not to compute the total damage consisting of fatalities and injuries (fatalities + injuries) since they have a different order of magnitude (a damage related to 1 death is far greater than a damage related to a light injury, for example). Throughout this report, I have always presented the data relating to fatalities and injuries separately.

Question 02 : Across the United States, which types of events have the greatest economic consequences?

*Property and Crops combined Economic Damage As for the impact on economy, we have got two sorted lists below by the amount of money cost by damages.

Aggregate Data for Property Damage.

```
propdmg <- aggregate(PROPDMG ~ CLEANNAME, data = cleandata, FUN = sum)
propdmg <- propdmg[order(propdmg$PROPDMG, decreasing = TRUE), ]
# 5 most harmful causes of injuries
propdmgMax <- propdmg[1:10, ]
print(propdmgMax)
```

```
##             CLEANNAME   PROPDMG
## 2               Flood 682475269
## 1         Flash Flood 681884269
## 8  Thunderstorm Wind 208637243
## 9             Tornado  10220134
## 3                Hail   3000236
## 7           Lightning   1720147
## 4          Heavy Snow    170000
## 5           High Wind        80
## 6           Ice Storm        50
## 10      Winter Storm         1
```

Aggregate crops damaged.

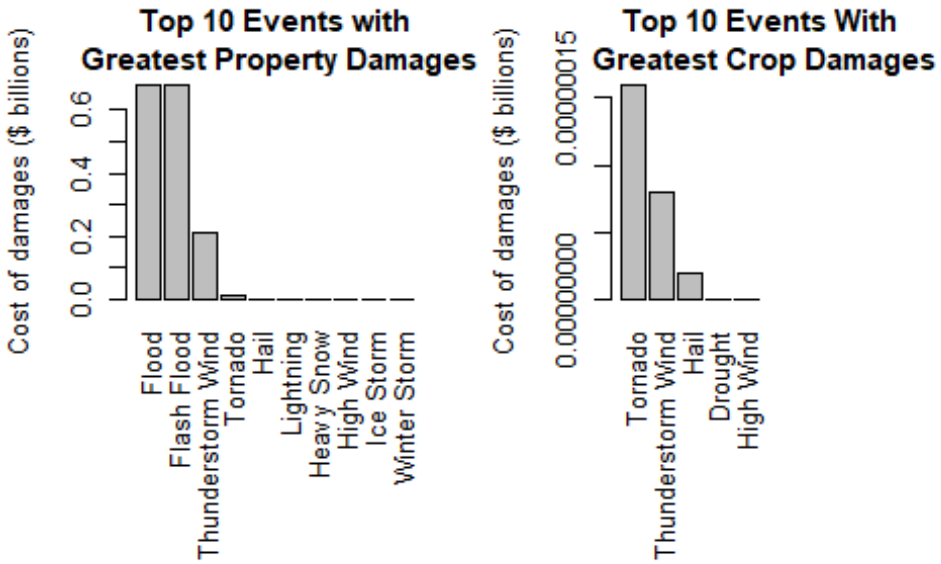
```
cropdmg <- aggregate(CROPDMG ~ CLEANNAME, data = cleandata, FUN = sum)
cropdmg <- cropdmg[order(cropdmg$CROPDMG, decreasing = TRUE), ]
# 5 most harmful causes of injuries
cropdmgMax <- cropdmg[1:10, ]
print(cropdmgMax)
```

```
##              CLEANNAME CROPDMG
## 5              Tornado     160
## 4    Thunderstorm Wind      80
## 2                 Hail      20
## 1              Drought       0
## 3            High Wind       0
## NA                <NA>      NA
## NA.1              <NA>      NA
## NA.2              <NA>      NA
## NA.3              <NA>      NA
## NA.4              <NA>      NA
```

Aggregate economy damage.

Plot showing severe effects on property,crops and the economy.

```
par(mfrow = c(1, 2), mar = c(15, 4, 3, 2), mgp = c(3, 1, 0), cex = 0.8)
barplot(propdmgMax$PROPDMG/(10^9), las = 3, names.arg = propdmgMax$CLEANNAME,
main = "Top 10 Events with\n Greatest Property Damages", ylab = "Cost of
damages ($ billions)", col = "grey")
barplot(cropdmgMax$CROPDMG/(10^9), las = 3, names.arg = cropdmgMax$CLEANNAME,
main = "Top 10 Events With\n Greatest Crop Damages", ylab = "Cost of damages
($ billions)", col = "grey")
```



The weather events have the Greatest Economic Consequences are: Flood, Drought, Tornado and Typhoon. Across the United States, Flood, Tornado and Typhoon have caused the Greatest Damage to Properties. *Drought and Flood had been the causes for the Greatest Damage to Crops.

Conclusion

From these data, we found that *Excessive Heat* and *Tornado* are most harmful with respect to Population Health, while *Flood*, *Drought* and *Hurricane/Typhoon* have the greatest Economic Consequences.