US Storms and its impact on Life and Economy

Ajith Samraj

Introduction

This is the second course project in the course Reproducible Research in Coursera. Human life and Property were damaged occasionally by natural disasters such as Storms , Tsunami etc., In this analysis we focus on Storm and its impact in Life and Property.

This study is based on the analysis of the weather conditions in U.S. over a period. The data is given by National Oceanic and Atmospheric Administration (NOAA). It includes all the data such as damage to Life as well as Property.

For more details regarding the data visit https://www.ncdc.noaa.gov/stormevents/

Synopsis

Here we are going to figure out answers for two questions

1 - which types of events cause most harm to population health?

2 - which types of events affects the economy?

Loading Libraries

```
library(knitr)
library(ggplot2)
library(dplyr)
library(plyr)
library(stats)
library(rmarkdown)
library(tinytex)
```

Loading Data

```
#Loading the Data
RawData <- read.csv("StormData.csv.bz2")
dim(RawData)
```

```
## [1] 902297     37
```

```
names(RawData)
```

```
##  [1] "STATE__"    "BGN_DATE"   "BGN_TIME"   "TIME_ZONE"  "COUNTY"
##  [6] "COUNTYNAME" "STATE"      "EVTYPE"     "BGN_RANGE"  "BGN_AZI"
## [11] "BGN_LOCATI" "END_DATE"   "END_TIME"   "COUNTY_END" "COUNTYENDN"
## [16] "END_RANGE"  "END_AZI"    "END_LOCATI" "LENGTH"     "WIDTH"
## [21] "F"          "MAG"        "FATALITIES" "INJURIES"   "PROPDMG"
## [26] "PROPDMGEXP" "CROPDMG"    "CROPDMGEXP" "WFO"        "STATEOFFIC"
## [31] "ZONENAMES"  "LATITUDE"   "LONGITUDE"  "LATITUDE_E" "LONGITUDE_"
## [36] "REMARKS"    "REFNUM"
```

There are 902297 observations with 37 variables. We need only a subset of the whole data for analysis.

We need only the following variables for analysing

*EVTYPE: Event Type

*FATALITIES: No of Fatalities

*INJURIES: No of Injuries

*PROGDMG: Property Damage

*PROPDMGEXP: Units for Property Damage (magnitudes - K,B,M)

*CROPDMG: Crop Damage

*CROPDMGEXP: Units for Crop Damage (magnitudes - K,BM,B)

```
#Selecting the Needed variables
vars <- c("EVTYPE", "FATALITIES", "INJURIES", "PROPDMG", "PROPDMGEXP", "CROPDMG", "CROPDMGEXP")
RawData <- RawData[vars]
dim(RawData)
```

```
## [1] 902297      7
```

Property Damage

```
unique(RawData$PROPDMGEXP)
```

```
##  [1] "K" "M" ""  "B" "m" "+" "0" "5" "6" "?" "4" "2" "3" "h" "7" "H" "-" "1" "8"
```

Some of the value of the column PROPDMGEXP are in Letters and Some in Numbers. We need to Covert them to a single unit.

```
RawData$PROPDMGEXP <- plyr::mapvalues(RawData$PROPDMGEXP, from = c("K", "M","", "B", "m", "+", "0", "5"
RawData$PROPDMGEXP <- as.numeric(as.character(RawData$PROPDMGEXP))
RawData$PROPDMGTOTAL <- (RawData$PROPDMG * RawData$PROPDMGEXP)/1000000000
```

Units of Crop Damage

```
unique(RawData$CROPDMGEXP)
```

```
## [1] ""  "M" "K" "m" "B" "?" "0" "k" "2"
```

Like above some value of the column CROPDMGEXP are in Letters and Some in Numbers. So we do conversion.

```
RawData$CROPDMGEXP <- mapvalues(RawData$CROPDMGEXP, from = c("","M", "K", "m", "B", "?", "0", "k","2"),
                      to = c(1,10^6, 10^3, 10^6, 10^9, 0, 1, 10^3, 10^2))
RawData$CROPDMGEXP <- as.numeric(as.character(RawData$CROPDMGEXP))
RawData$CROPDMGTOTAL <- (RawData$CROPDMG * RawData$CROPDMGEXP)/1000000000
```

Events with more Fatality rate

Fatality rate

Analysing which events cause most Fatalities . FATILITIES is the factor variable. Almost 985 events were recorded by NOAA.

```
sumFatalities <- aggregate(FATALITIES ~ EVTYPE, data = RawData,  FUN="sum")
dim(sumFatalities)
```

```
## [1] 985   2
```

Looking top 5 fatal events.

```
fatalities5events <- sumFatalities[order(-sumFatalities$FATALITIES), ][1:5, ]
dim(fatalities5events)
```

```
## [1] 5 2
```

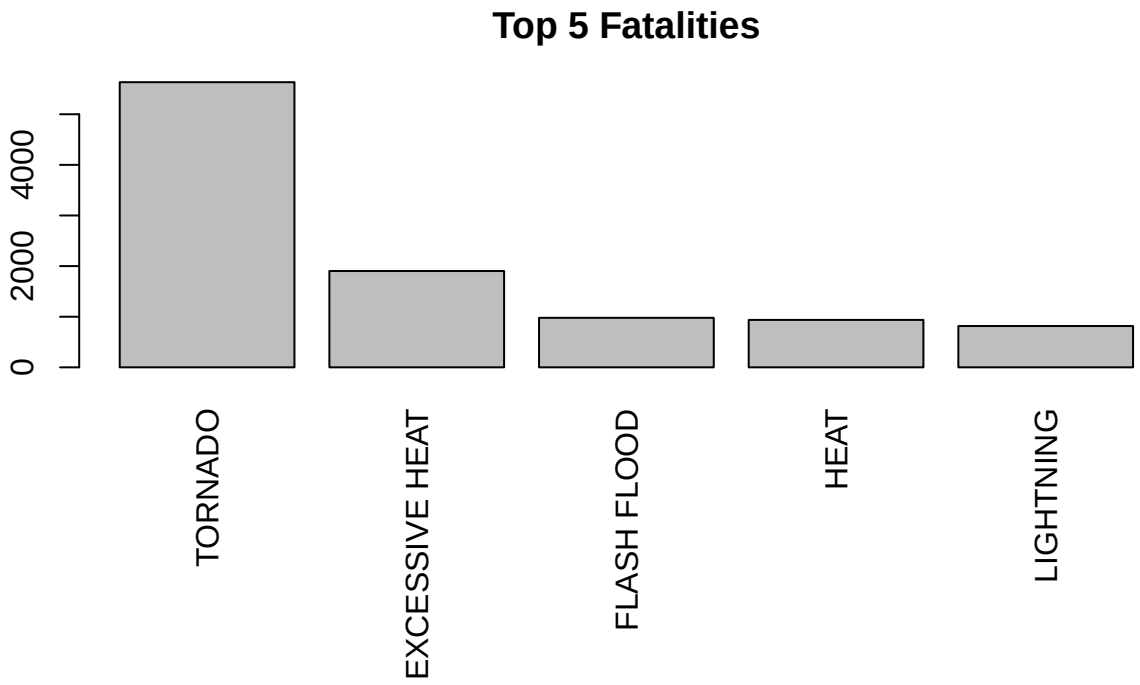
```
fatalities5events
```

```
##             EVTYPE FATALITIES
## 834        TORNADO       5633
## 130 EXCESSIVE HEAT       1903
## 153    FLASH FLOOD        978
## 275           HEAT        937
## 464      LIGHTNING        816
```

Fatality Plot

Plot of the Top 5 fatal events.

```
par(mar=c(12, 3, 3, 1))
barplot(fatalities5events$FATALITIES, names.arg = fatalities5events$EVTYPE, las = 3,
        main = "Top 5 Fatalities", ylab = "No.of Fatalities")
```



Injury rate

Events which are causing injuries to human life.

```
sumInjuries <- aggregate(INJURIES ~ EVTYPE, data = RawData,  FUN="sum")
dim(sumInjuries)
```

```
## [1] 985   2
```

Sorting Top 5 injury causing events.

```
injuries5events <- sumInjuries[order(-sumInjuries$INJURIES), ][1:5, ]
dim(injuries5events)
```

```
## [1] 5 2
```

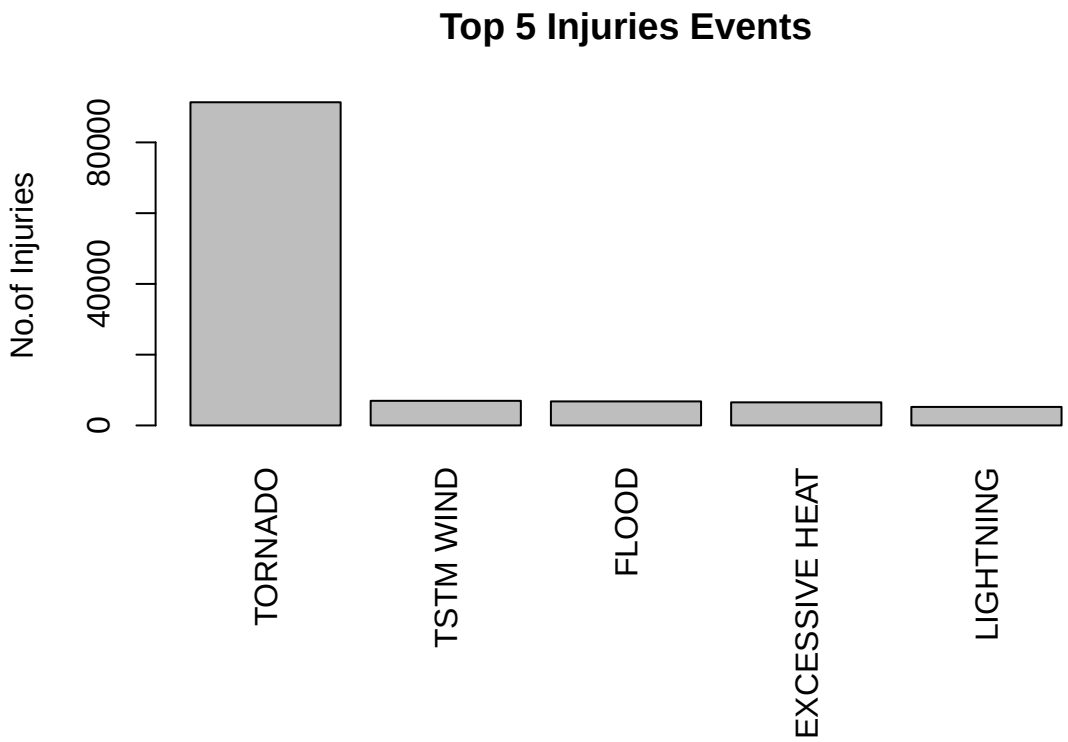
Top 5 Injury events

```
injuries5events
```

```
##             EVTYPE INJURIES
## 834        TORNADO    91346
## 856      TSTM WIND     6957
## 170          FLOOD     6789
## 130 EXCESSIVE HEAT     6525
## 464      LIGHTNING     5230
```

Plot of the top 5 injury causing events

```
par(mar=c(10,4,4,4))
barplot(injuries5events$INJURIES, names.arg = injuries5events$EVTYPE, las = 3,
        main = "Top 5 Injuries Events", ylab = "No.of Injuries")
```



Events with more Economic damage

We have to rely on PROPDMG (Property Damage) and CROPDMG (Crop Damage) to figue out the Economic impacts.

Crop Damage

Crop damage is given by the variable CROPDMG.

```
sumCropDamage <- aggregate(CROPDMGTOTAL ~ EVTYPE, data = RawData,  FUN="sum")
dim(sumCropDamage)
```

```
## [1] 985   2
```

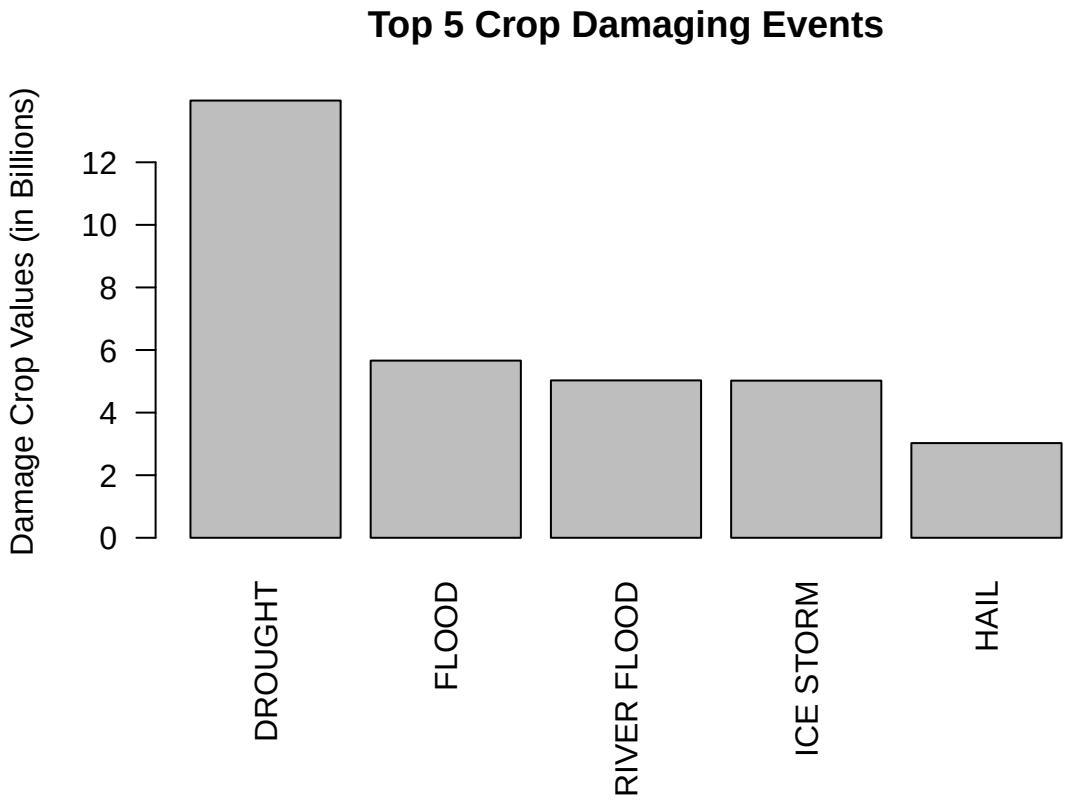
Top 5 Crop damaging events

```
cropdmg5Total <- sumCropDamage[order(-sumCropDamage$CROPDMGTOTAL), ][1:5, ]
cropdmg5Total
```

```
##           EVTYPE CROPDMGTOTAL
## 95       DROUGHT    13.972566
## 170        FLOOD     5.661968
## 590 RIVER FLOOD     5.029459
## 427   ICE STORM     5.022113
## 244         HAIL     3.025954
```

Plot of the Top 5 Crop damaging events

```
par(mar=c(7,4,4,4))
barplot(cropdmg5Total$CROPDMGTOTAL, names.arg = cropdmg5Total$EVTYPE, las = 2,
        main = "Top 5 Crop Damaging Events",
        ylab = "Damage Crop Values (in Billions)")
```



Property Damage Events

```
sumPropertyDamage <- aggregate(PROPDMGTOTAL ~ EVTYPE, data = RawData,  FUN="sum")
dim(sumPropertyDamage)
```

```
## [1] 985   2
```

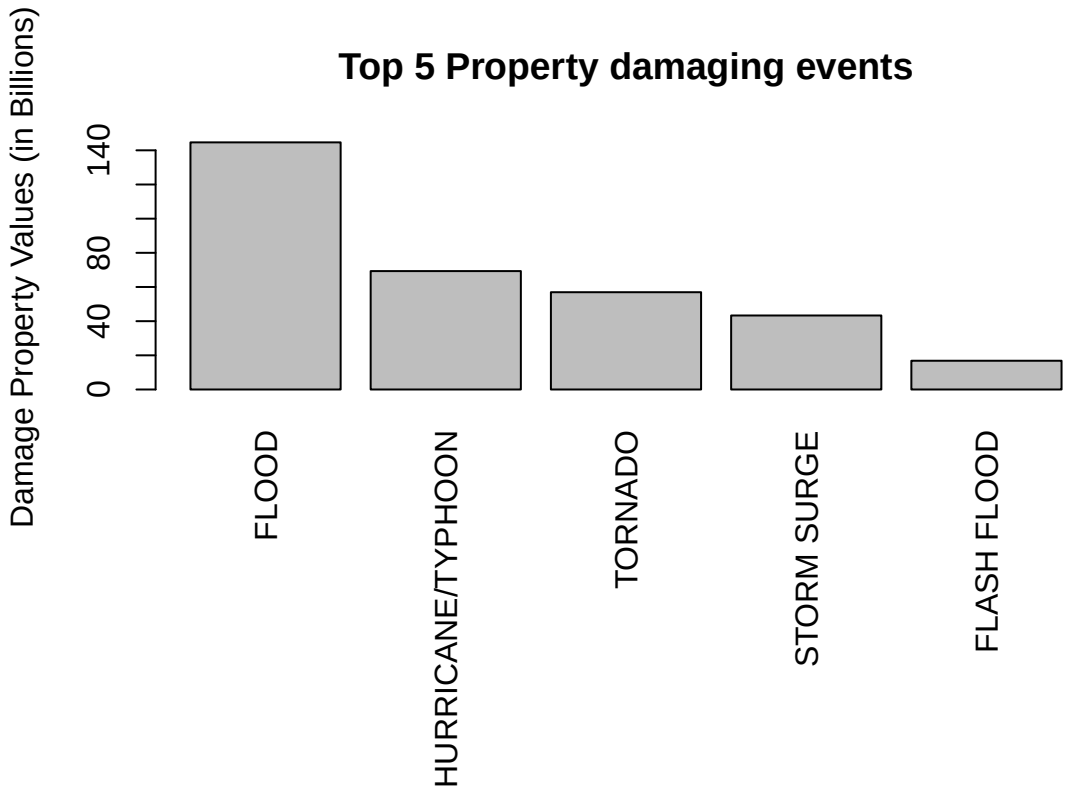
Top 5 Property damage events

```
propdmg5Total <- sumPropertyDamage[order(-sumPropertyDamage$PROPDMGTOTAL), ][1:5, ]
propdmg5Total
```

```
##                EVTYPE PROPDMGTOTAL
## 170             FLOOD    144.65771
## 411 HURRICANE/TYPHOON     69.30584
## 834           TORNADO     56.94738
## 670       STORM SURGE     43.32354
## 153       FLASH FLOOD     16.82267
```

Plot for Top 5 property Damaging events

```
par(mar=c(12,4,4,4))
barplot(propdmg5Total$PROPDMGTOTAL, names.arg = propdmg5Total$EVTYPE, las = 3,
        main = "Top 5 Property damaging events",
        ylab = "Damage Property Values (in Billions)")
```



Results

Question - 1

Which Event causes more Fatalities ?

The results tells us that **Tornados** causes the highest number of Fatalities and Injuries.

Question - 2

Which Event causes more Economic consequences ?

Since we analaysed on two perspectives we got two results according to the property.

The results tells us that **Floods** causes highest **Property Damage**.

The results tells us that **Droughts** causes highest **Crop damages**.