

Damage due to Weather Events

Himank Jain

March 15, 2019

Synopsis: Analysis of Weather Events.

The source of the data is NOAA Storm Database. The Analysis tries to deduce the effect of multiple weather events on Public Health and Economy. Weather events cause massive health hazards and economic problems every year and it's important to understand their nature and characteristics. This analysis involves data visualization for studying the effects of major weather events on both public health and economy. The health effects include fatalities and injuries and economic factors include property damage and crop damage.

Data Processing:

Loading required Packages

Reading the data

```
storm<-fread("repdata_data_StormData.csv.bz2",sep="," ,header=TRUE)
```

Exploratory Analysis

```
storm<-tbl_df(storm)
summary(storm)
```

```
##      STATE__      BGN_DATE      BGN_TIME      TIME_ZONE
##  Min.   : 1.0    Length:902297    Length:902297    Length:902297
##  1st Qu.:19.0    Class :character    Class :character    Class :character
##  Median :30.0    Mode  :character    Mode  :character    Mode  :character
##  Mean   :31.2
##  3rd Qu.:45.0
##  Max.   :95.0
##
##      COUNTY      COUNTYNAME      STATE      EVTYPE
##  Min.   : 0.0    Length:902297    Length:902297    Length:902297
##  1st Qu.:31.0    Class :character    Class :character    Class :character
##  Median :75.0    Mode  :character    Mode  :character    Mode  :character
##  Mean   :100.6
##  3rd Qu.:131.0
##  Max.   :873.0
##
##      BGN_RANGE      BGN_AZI      BGN_LOCATI      END_DATE
##  Min.   : 0.000    Length:902297    Length:902297    Length:902297
##  1st Qu.: 0.000    Class :character    Class :character    Class :character
```

```

## Median : 0.000 Mode :character Mode :character Mode :character
## Mean : 1.484
## 3rd Qu.: 1.000
## Max. :3749.000
##
## END_TIME COUNTY_END COUNTYENDN END_RANGE
## Length:902297 Min. :0 Mode:logical Min. : 0.0000
## Class :character 1st Qu.:0 NA's:902297 1st Qu.: 0.0000
## Mode :character Median :0 Median : 0.0000
## Mean :0 Mean : 0.9862
## 3rd Qu.:0 3rd Qu.: 0.0000
## Max. :0 Max. :925.0000
##
## END_AZI END_LOCATI LENGTH WIDTH
## Length:902297 Length:902297 Min. : 0.0000 Min. : 0.000
## Class :character Class :character 1st Qu.: 0.0000 1st Qu.: 0.000
## Mode :character Mode :character Median : 0.0000 Median : 0.000
## Mean : 0.2301 Mean : 7.503
## 3rd Qu.: 0.0000 3rd Qu.: 0.000
## Max. :2315.0000 Max. :4400.000
##
## F MAG FATALITIES INJURIES
## Min. :0.0 Min. : 0.0 Min. : 0.0000 Min. : 0.0000
## 1st Qu.:0.0 1st Qu.: 0.0 1st Qu.: 0.0000 1st Qu.: 0.0000
## Median :1.0 Median : 50.0 Median : 0.0000 Median : 0.0000
## Mean :0.9 Mean : 46.9 Mean : 0.0168 Mean : 0.1557
## 3rd Qu.:1.0 3rd Qu.: 75.0 3rd Qu.: 0.0000 3rd Qu.: 0.0000
## Max. :5.0 Max. :22000.0 Max. :583.0000 Max. :1700.0000
## NA's :843563
## PROPDMG PROPDMGEXP CROPDMG CROPDMGEXP
## Min. : 0.00 Length:902297 Min. : 0.000 Length:902297
## 1st Qu.: 0.00 Class :character 1st Qu.: 0.000 Class :character
## Median : 0.00 Mode :character Median : 0.000 Mode :character
## Mean : 12.06 Mean : 1.527
## 3rd Qu.: 0.50 3rd Qu.: 0.000
## Max. :5000.00 Max. :990.000
##
## WFO STATEOFFIC ZONENAMES LATITUDE
## Length:902297 Length:902297 Length:902297 Min. : 0
## Class :character Class :character Class :character 1st Qu.:2802
## Mode :character Mode :character Mode :character Median :3540
## Mean :2875
## 3rd Qu.:4019
## Max. :9706
## NA's :47
## LONGITUDE LATITUDE_E LONGITUDE_ REMARKS
## Min. : -14451 Min. : 0 Min. : -14455 Length:902297
## 1st Qu.: 7247 1st Qu.: 0 1st Qu.: 0 Class :character
## Median : 8707 Median : 0 Median : 0 Mode :character
## Mean : 6940 Mean :1452 Mean : 3509
## 3rd Qu.: 9605 3rd Qu.:3549 3rd Qu.: 8735
## Max. : 17124 Max. :9706 Max. :106220
## NA's :40
## REFNUM

```

```
## Min.      :    1
## 1st Qu.:225575
## Median :451149
## Mean      :451149
## 3rd Qu.:676723
## Max.      :902297
##
```

```
str(storm)
```

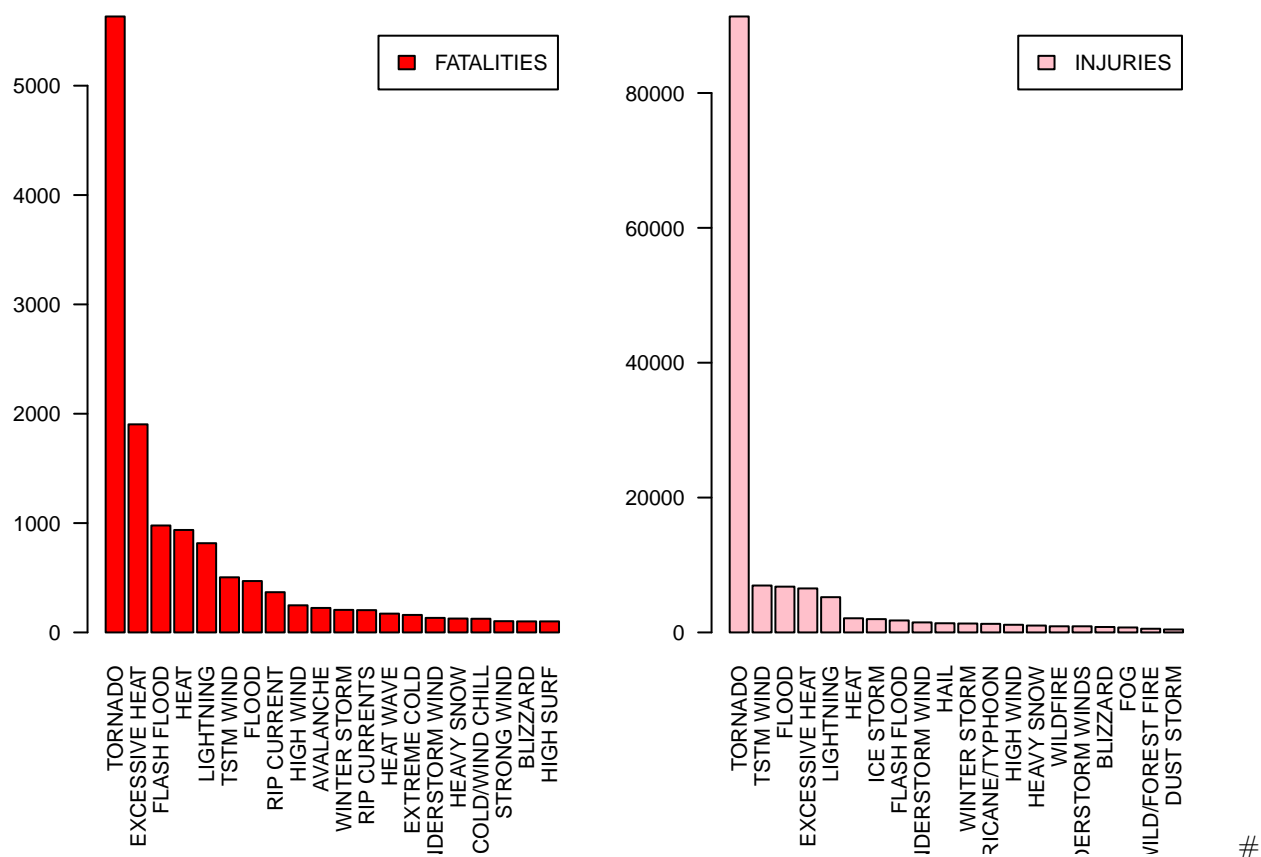
```
## Classes 'tbl_df', 'tbl' and 'data.frame':   902297 obs. of  37 variables:
## $ STATE__      : num  1 1 1 1 1 1 1 1 1 1 ...
## $ BGN_DATE     : chr   "4/18/1950 0:00:00" "4/18/1950 0:00:00" "2/20/1951 0:00:00" "6/8/1951 0:00:00" .
## $ BGN_TIME     : chr   "0130" "0145" "1600" "0900" ...
## $ TIME_ZONE    : chr   "CST" "CST" "CST" "CST" ...
## $ COUNTY       : num  97 3 57 89 43 77 9 123 125 57 ...
## $ COUNTYNAME   : chr   "MOBILE" "BALDWIN" "FAYETTE" "MADISON" ...
## $ STATE        : chr   "AL" "AL" "AL" "AL" ...
## $ EVTYPE       : chr   "TORNADO" "TORNADO" "TORNADO" "TORNADO" ...
## $ BGN_RANGE    : num  0 0 0 0 0 0 0 0 0 0 ...
## $ BGN_AZI      : chr   "" "" "" "" ...
## $ BGN_LOCATI   : chr   "" "" "" "" ...
## $ END_DATE     : chr   "" "" "" "" ...
## $ END_TIME     : chr   "" "" "" "" ...
## $ COUNTY_END   : num  0 0 0 0 0 0 0 0 0 0 ...
## $ COUNTYENDN   : logi  NA NA NA NA NA NA ...
## $ END_RANGE    : num  0 0 0 0 0 0 0 0 0 0 ...
## $ END_AZI      : chr   "" "" "" "" ...
## $ END_LOCATI   : chr   "" "" "" "" ...
## $ LENGTH       : num  14 2 0.1 0 0 1.5 1.5 0 3.3 2.3 ...
## $ WIDTH        : num  100 150 123 100 150 177 33 33 100 100 ...
## $ F            : int   3 2 2 2 2 2 2 1 3 3 ...
## $ MAG          : num  0 0 0 0 0 0 0 0 0 0 ...
## $ FATALITIES   : num  0 0 0 0 0 0 0 0 1 0 ...
## $ INJURIES     : num  15 0 2 2 2 2 6 1 0 14 0 ...
## $ PROPDMG      : num  25 2.5 25 2.5 2.5 2.5 2.5 2.5 25 25 ...
## $ PROPDMGEXP   : chr   "K" "K" "K" "K" ...
## $ CROPDMG      : num  0 0 0 0 0 0 0 0 0 0 ...
## $ CROPDMGEXP   : chr   "" "" "" "" ...
## $ WFO          : chr   "" "" "" "" ...
## $ STATEOFFIC   : chr   "" "" "" "" ...
## $ ZONENAMES    : chr   "" "" "" "" ...
## $ LATITUDE     : num  3040 3042 3340 3458 3412 ...
## $ LONGITUDE    : num  8812 8755 8742 8626 8642 ...
## $ LATITUDE_E   : num  3051 0 0 0 0 ...
## $ LONGITUDE_   : num  8806 0 0 0 0 ...
## $ REMARKS      : chr   "" "" "" "" ...
## $ REFNUM       : num  1 2 3 4 5 6 7 8 9 10 ...
## - attr(*, ".internal.selfref")=<externalptr>
```

1) PUBLIC HEALTH:Fatalities and Injuries

```
stormpop<-subset(storm,select = c("EVTYPE","FATALITIES","INJURIES"))
fatalities<-aggregate(data=stormpop,FATALITIES~EVTYPE,FUN = sum)
fatalities<-fatalities[order(fatalities$FATALITIES,decreasing = TRUE),]
injuries<-aggregate(data=stormpop,INJURIES~EVTYPE,FUN=sum)
injuries<-injuries[order(injuries$INJURIES,decreasing = TRUE),]
topfatalities<-fatalities[1:20,]
topinjuries<-injuries[1:20,]
```

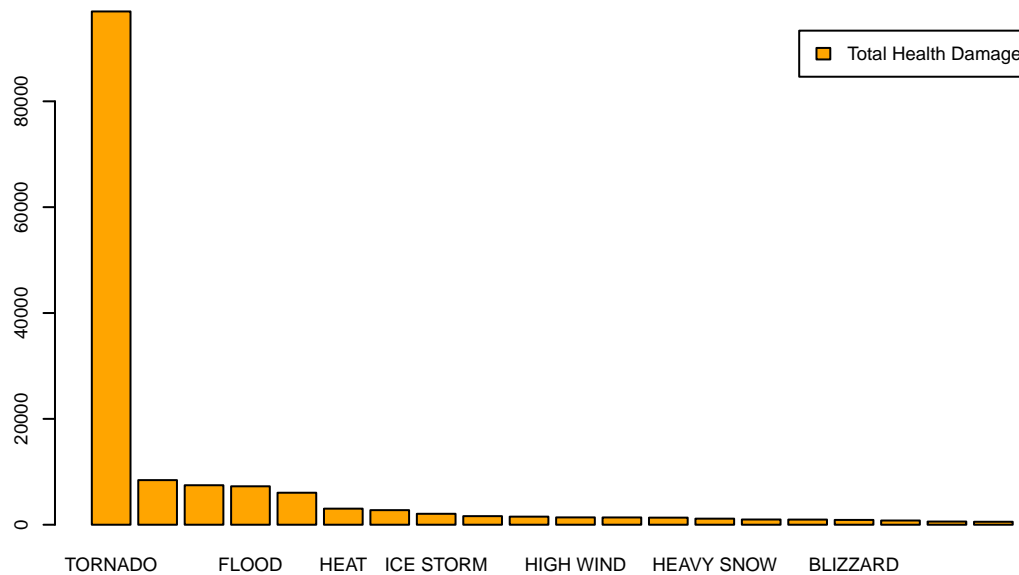
##Plotting Most effective Weather Events vs Fatalities and Most effective Weather Events and Injuries.

```
par(mfrow = c(1, 2), las = 2,cex=0.7,font.lab=2,mar=c(8,4,1,1))
barplot(topfatalities$FATALITIES,names.arg=topfatalities$EVTYPE,col="red",legend.text = "FATALITIES")
barplot(topinjuries$INJURIES,names.arg = topinjuries$EVTYPE,col="pink",legend.text = "INJURIES")
```



Aggregating Fatalities and Injuries:

```
healthdmg<-merge(fatalities,injuries,by="EVTYPE")
healthdmg$dmg<-healthdmg$FATALITIES+healthdmg$INJURIES
healthdmg<-healthdmg[order(healthdmg$dmg,decreasing = TRUE),]
tophealthdmg<-healthdmg[1:20,]
par(mfrow=c(1,1),mar=c(12,8,3,3),cex=0.6)
barplot(tophealthdmg$dmg,names.arg = tophealthdmg$EVTYPE,col="orange",legend.text = "Total Health Damage")
```



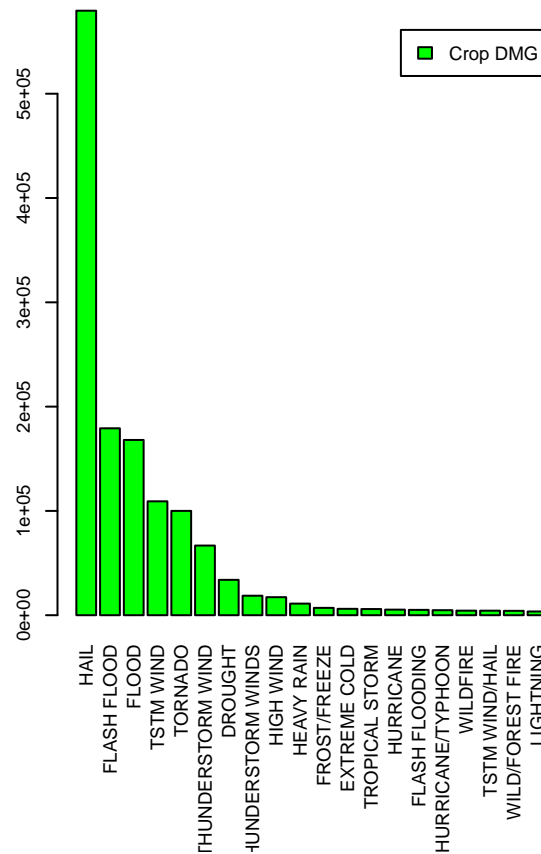
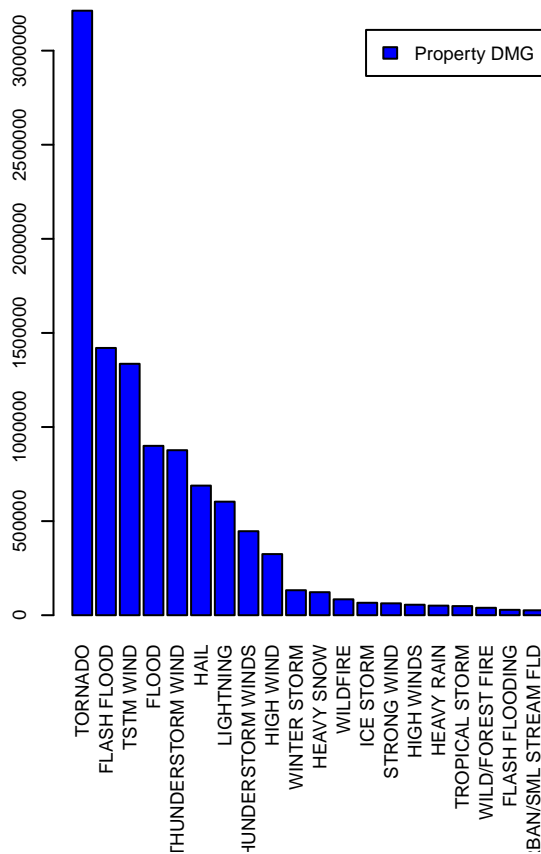
- We can Observe that most of the Public health effects (fatalities and injuries) are caused by Tornadoes.

2) Economic Consequences:

```
stormprop<-subset(storm,select=c("EVTYPE", "PROPDMG", "CROPDMG"))
propdmg<-aggregate(data=stormprop,PROPDMG~EVTYPE,FUN=sum)
croppdmg<-aggregate(data=stormprop,CROPDMG~EVTYPE,FUN=sum)
propdmg<-propdmg[order(propdmg$PROPDMG,decreasing = TRUE),]
croppdmg<-croppdmg[order(croppdmg$CROPDMG,decreasing=TRUE),]
topprop<-propdmg[1:20,]
topcrop<-croppdmg[1:20,]
```

Visualizing Significant Weather Events vs Property Damage and Significant Weather Events vs Crop Damage:

```
par(mfrow = c(1, 2), las = 3,cex=0.6,font.lab=2,mar=c(10,4,1,1))
barplot(topprop$PROPDMG,names.arg = topprop$EVTYPE,col="blue",legend.text = "Property DMG")
barplot(topcrop$CROPDMG,names.arg = topcrop$EVTYPE,col="green",legend.text = "Crop DMG")
```



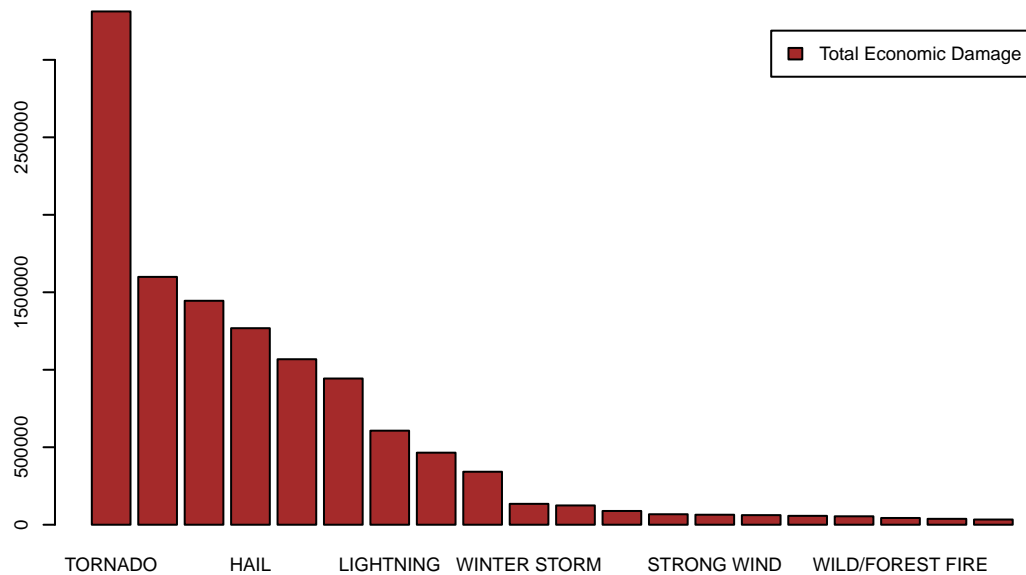
##

Aggregating Property and Crop Damages:

```
economicdmg<-merge(propdmg,cropdmg,by="EVTYPE")
economicdmg$ecodmg<- economicdmg$PROPDMG+economicdmg$CROPDMG
economicdmg<-economicdmg[order(economicdmg$ecodmg,decreasing = TRUE),]
topecodmg<-economicdmg[1:20,]
```

Significant Weather Events vs Economic Loss

```
par(mfrow=c(1,1),mar=c(12,8,3,3),cex=0.6)
barplot(topecodmg$ecodmg,names.arg=topecodmg$EVTYPE,col="brown",legend.text = "Total Economic Damage")
```



* We can observe that most of the Economic loss was due to, yet again, Tornadoes!

Results:

From the Analysis done above on the NOAA Storm Database it is quite clear that there are huge health effects and economic consequences due to multiple Weather events in USA alone. Most of these Harmful Health Hazards and Economic loss can be associated largely with **Tornadoes**. There are many other weather events which also harm population health and cause economic losses. Second most deaths are caused by **Excessive Heat** and Second most injuries by **TSTM WINDS**. Second most Property Damage and Crop Damage is done by **Flash Floods**.