Most Damaging Severe Weather Across USA v2

Synopsis

By analysing the NOAA Storm database, and cleanning up the data extensively by removing inconsistencies, we have reached the following conclusions:

1-The most harmful severe weather to the human health is tornados

2-The most damaging severe weather to human properties and to the crops is floods.

Preprocessing Data

Loading required libraries

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.1.2
```

```
library(graphics)
library(xtable)
```

Downloading the data

```
fileUrl = "https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2"
#download.file(fileUrl, destfile = "..//data//repdata_data_StormData.csv.bz2", method = "auto")
file.bz <- bzfile("..//data//StormData.zip","rt")

stormDataDownloaded = date()
```

Loading the data

```
#stormData = read.csv("..//data//repdata_data_StormData.csv", sep = ",")

stormData = read.csv(file.bz)
```

Summary of Storm Data

```
summary(stormData)
```

```
##     STATE__                  BGN_DATE             BGN_TIME
##  Min.   : 1.0   5/25/2011 0:00:00:  1202   12:00:00 AM: 10163
##  1st Qu.:19.0   4/27/2011 0:00:00:  1193   06:00:00 PM:  7350
##  Median :30.0   6/9/2011 0:00:00 :  1030   04:00:00 PM:  7261
##  Mean   :31.2   5/30/2004 0:00:00:  1016   05:00:00 PM:  6891
##  3rd Qu.:45.0   4/4/2011 0:00:00 :  1009   12:00:00 PM:  6703
##  Max.   :95.0   4/2/2006 0:00:00 :   981   03:00:00 PM:  6700
##                 (Other)          :895866   (Other)    :857229
##    TIME_ZONE          COUNTY         COUNTYNAME         STATE
##  CST    :547493   Min.   :  0   JEFFERSON :  7840   TX     : 83728
##  EST    :245558   1st Qu.: 31   WASHINGTON:  7603   KS     : 53440
##  MST    : 68390   Median : 75   JACKSON   :  6660   OK     : 46802
##  PST    : 28302   Mean   :101   FRANKLIN  :  6256   MO     : 35648
##  AST    :  6360   3rd Qu.:131   LINCOLN   :  5937   IA     : 31069
##  HST    :  2563   Max.   :873   MADISON   :  5632   NE     : 30271
##  (Other):  3631                 (Other)   :862369   (Other):621339
##                EVTYPE         BGN_RANGE       BGN_AZI
##  HAIL              :288661   Min.   :   0          :547332
##  TSTM WIND         :219940   1st Qu.:   0   N      : 86752
##  THUNDERSTORM WIND: 82563   Median :   0   W      : 38446
##  TORNADO           : 60652   Mean   :   1   S      : 37558
##  FLASH FLOOD       : 54277   3rd Qu.:   1   E      : 33178
##  FLOOD             : 25326   Max.   :3749   NW     : 24041
##  (Other)           :170878                  (Other):134990
##          BGN_LOCATI                  END_DATE             END_TIME
##                :287743                     :243411             :238978
##  COUNTYWIDE    : 19680   4/27/2011 0:00:00:  1214   06:00:00 PM:  9802
##  Countywide    :   993   5/25/2011 0:00:00:  1196   05:00:00 PM:  8314
##  SPRINGFIELD   :   843   6/9/2011 0:00:00 :  1021   04:00:00 PM:  8104
##  SOUTH PORTION:   810   4/4/2011 0:00:00 :  1007   12:00:00 PM:  7483
##  NORTH PORTION:   784   5/30/2004 0:00:00:   998   11:59:00 PM:  7184
##  (Other)       :591444   (Other)          :653450   (Other)    :622432
##    COUNTY_END COUNTYENDN       END_RANGE       END_AZI
##  Min.   :0    Mode:logical   Min.   :  0          :724837
##  1st Qu.:0    NA's:902297    1st Qu.:  0   N      : 28082
##  Median :0                   Median :  0   S      : 22510
##  Mean   :0                   Mean   :  1   W      : 20119
##  3rd Qu.:0                   3rd Qu.:  0   E      : 20047
##  Max.   :0                   Max.   :925   NE     : 14606
##                                            (Other): 72096
##            END_LOCATI         LENGTH           WIDTH            F
##                 :499225   Min.   :   0.0   Min.   :   0   Min.   :0
##  COUNTYWIDE     : 19731   1st Qu.:   0.0   1st Qu.:   0   1st Qu.:0
##  SOUTH PORTION  :   833   Median :   0.0   Median :   0   Median :1
##  NORTH PORTION  :   780   Mean   :   0.2   Mean   :   8   Mean   :1
##  CENTRAL PORTION:   617   3rd Qu.:   0.0   3rd Qu.:   0   3rd Qu.:1
##  SPRINGFIELD    :   575   Max.   :2315.0   Max.   :4400   Max.   :5
##  (Other)        :380536                                  NA's   :843563
##       MAG          FATALITIES     INJURIES         PROPDMG
##  Min.   :    0   Min.   :  0   Min.   :   0.0   Min.   :   0
##  1st Qu.:    0   1st Qu.:  0   1st Qu.:   0.0   1st Qu.:   0
##  Median :   50   Median :  0   Median :   0.0   Median :   0
##  Mean   :   47   Mean   :  0   Mean   :   0.2   Mean   :  12
##  3rd Qu.:   75   3rd Qu.:  0   3rd Qu.:   0.0   3rd Qu.:   0
```

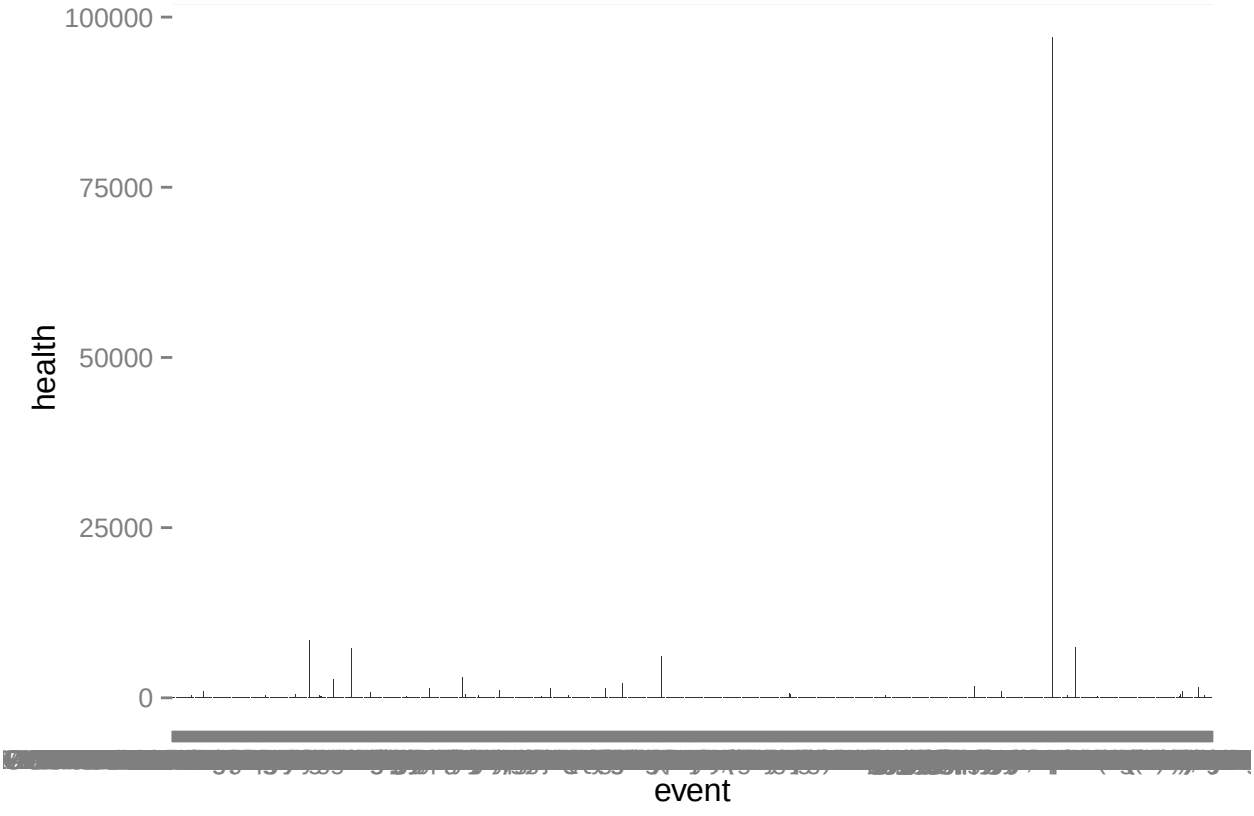
```
##  Max.   :22000   Max.   :583   Max.   :1700.0   Max.   :5000
##
##    PROPDMGEXP        CROPDMG        CROPDMGEXP          WFO
##         :465934   Min.   :  0.0          :618413          :142069
##  K      :424665   1st Qu.:  0.0   K      :281832   OUN    : 17393
##  M      : 11330   Median :  0.0   M      :  1994   JAN    : 13889
##  0      :   216   Mean   :  1.5   k      :    21   LWX    : 13174
##  B      :    40   3rd Qu.:  0.0   0      :    19   PHI    : 12551
##  5      :    28   Max.   :990.0   B      :     9   TSA    : 12483
##  (Other):    84                   (Other):     9   (Other):690738
##                                STATEOFFIC
##                                     :248769
##  TEXAS, North                       : 12193
##  ARKANSAS, Central and North Central: 11738
##  IOWA, Central                      : 11345
##  KANSAS, Southwest                  : 11212
##  GEORGIA, North and Central         : 11120
##  (Other)                            :595920
##
##
##
##  GREATER RENO / CARSON CITY / M - GREATER RENO / CARSON CITY / M
##  GREATER LAKE TAHOE AREA - GREATER LAKE TAHOE AREA
##  JEFFERSON - JEFFERSON
##  MADISON - MADISON
##  (Other)
##     LATITUDE      LONGITUDE        LATITUDE_E     LONGITUDE_
##  Min.   :   0   Min.   :-14451   Min.   :   0   Min.   :-14455
##  1st Qu.:2802   1st Qu.:  7247   1st Qu.:   0   1st Qu.:     0
##  Median :3540   Median :  8707   Median :   0   Median :     0
##  Mean   :2875   Mean   :  6940   Mean   :1452   Mean   :  3509
##  3rd Qu.:4019   3rd Qu.:  9605   3rd Qu.:3549   3rd Qu.:  8735
##  Max.   :9706   Max.   : 17124   Max.   :9706   Max.   :106220
##  NA's   :47                      NA's   :40
##                                          REMARKS           REFNUM
##                                              :287433   Min.   :     1
##                                              : 24013   1st Qu.:225575
##  Trees down.\n                               :  1110   Median :451149
##  Several trees were blown down.\n            :   568   Mean   :451149
##  Trees were downed.\n                        :   446   3rd Qu.:676723
##  Large trees and power lines were blown down.\n:   432   Max.   :902297
##  (Other)                                     :588295
```

Data Processing

```
# Taking the sum of fatalities for each event type
fatalities = tapply(stormData$FATALITIES, stormData$EVTYPE, sum)
# Taking the sum of injuries for each event type
injuries = tapply(stormData$INJURIES, stormData$EVTYPE, sum)
# Taking the sum of property damage for each event type
propertyDamage = tapply(stormData$PROPDMG, stormData$EVTYPE, sum)
```

```
# Taking the sum of crop damage for each event type
cropDamage = tapply(stormData$CROPDMG, stormData$EVTYPE, sum)

#creating a new dataframe to store all economic and health damages for each type
consequences = NULL
consequences  = data.frame(event = dimnames(fatalities), injuries, fatalities, propertyDamage, cropDamag

names(consequences) = c("event", "injuries", "fatalities", "propertyDamage", "cropDamage")

# A simple summation of all fatalities and injuries can be a measure of how harmful the event is
consequences$health = consequences$injuries + consequences$fatalities

healthPlot = ggplot(consequences, aes(x=event, y=health)) + geom_bar(stat="identity")

healthPlot
```

```
## Warning: position_stack requires constant width: output may be incorrect
```




```
# A simple summation of all property and crop damages can be a measure of how damaging the event is
consequences$economic = propertyDamage + cropDamage
```

```
sortedHealth=consequences[order(consequences$health, decreasing = TRUE),c(1,2,3,6)]
sortedEconomic=consequences[order(consequences$economic, decreasing = TRUE),c(1,4,5,7)]
#sortedEconomic = sort()
```

Results

The types of events that are most harmful to population health

TORNADO ###Its harm can lead to a max of: 9.6979 × 104 injuries and fatalities

The types of events that have the greatest economic consequences

TORNADO ###Its damage is maxed at: 3.3123 × 106

```
#print(xtable(consequences), type="latex")
```