Health and Economic Impact of Weather Events in the US

Based on Sefa Kilic work.

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.

Data FAQ available on site https://www.ncdc.noaa.gov/stormevents/faq.jsp

Synopsis

The analysis on the storm event database revealed that tornadoes are the most dangerous weather event to the population health. The second most dangerous event type is the excessive heat. The economic impact of weather events was also analyzed. Flash floods and thunderstorm winds caused billions of dollars in property damages between 1950 and 2011. The largest crop damage caused by drought, followed by flood and hails.

Data Processing

The data for this assignment was extracted from 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.

Load some library:

```
library(plyr)
library(ggplot2)
library(grid)
library(gridExtra)
```

```
rawdata <- read.csv(bzfile("repdata_data_StormData.csv.bz2"))
```

```
length(unique(rawdata$EVTYPE))
```

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

```
et <- tolower(rawdata$EVTYPE)
et <- gsub("[[:blank:][:punct:]+]", " ", et)

length(unique(et))
```

```
## [1] 874
```

```
rawdata$EVTYPE <- et
```

Across the United States, which types of events (as indicated in the EVTYPE variable) are most harmful with respect to population health?

```
reasons <- ddply(rawdata, .(EVTYPE), summarize,
                 fatalities = sum(FATALITIES),
                 injuries = sum(INJURIES))

fatal <- head(reasons[order(reasons$fatalities, decreasing = TRUE), ], 10)
injury <- head(reasons[order(reasons$injuries, decreasing = TRUE), ], 10)
```

```
fatal[, c("EVTYPE", "fatalities")]
```

```
##             EVTYPE fatalities
## 737        tornado       5633
## 109 excessive heat       1903
## 132    flash flood        978
## 234           heat        937
## 400      lightning        816
## 760      tstm wind        504
## 148          flood        470
## 511    rip current        368
## 309      high wind        248
## 11       avalanche        224
```

```
injury[, c("EVTYPE", "injuries")]
```

```
##                EVTYPE injuries
## 737           tornado    91346
## 760         tstm wind     6957
## 148             flood     6789
## 109    excessive heat     6525
## 400         lightning     5230
## 234              heat     2100
## 377         ice storm     1975
## 132       flash flood     1777
## 670 thunderstorm wind     1488
## 203              hail     1361
```

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

```
exp_transform <- function(e) {
    # h -> hundred, k -> thousand, m -> million, b -> billion
    if (e %in% c('h', 'H'))
        return(2)
    else if (e %in% c('k', 'K'))
```

```
        return(3)
    else if (e %in% c('m', 'M'))
        return(6)
    else if (e %in% c('b', 'B'))
        return(9)
    else if (!is.na(as.numeric(e))) # if a digit
        return(as.numeric(e))
    else if (e %in% c('', '-', '?', '+'))
        return(0)
    else {
        stop("Invalid exponent value.")
    }
}
```

```
prop_dmg_exp <- sapply(rawdata$PROPDMGEXP, FUN=exp_transform)
rawdata$prop_dmg <- rawdata$PROPDMG * (10 ** prop_dmg_exp)
crop_dmg_exp <- sapply(rawdata$CROPDMGEXP, FUN=exp_transform)
rawdata$crop_dmg <- rawdata$CROPDMG * (10 ** crop_dmg_exp)

econ_loss <- ddply(rawdata, .(EVTYPE), summarize,
                   prop_dmg = sum(prop_dmg),
                   crop_dmg = sum(crop_dmg))


econ_loss <- econ_loss[(econ_loss$prop_dmg > 0 | econ_loss$crop_dmg > 0), ]
prop_dmg_events <- head(econ_loss[order(econ_loss$prop_dmg, decreasing = TRUE), ], 10)
crop_dmg_events <- head(econ_loss[order(econ_loss$crop_dmg, decreasing = TRUE), ], 10)
```

```
prop_dmg_events[, c("EVTYPE", "prop_dmg")]
```

```
##                 EVTYPE     prop_dmg
## 132        flash flood 6.820237e+13
## 694 thunderstorm winds 2.086532e+13
## 737            tornado 1.078951e+12
## 203               hail 3.157558e+11
## 400          lightning 1.729433e+11
## 148              flood 1.446577e+11
## 361  hurricane typhoon 6.930584e+10
## 155           flooding 5.920826e+10
## 581        storm surge 4.332354e+10
## 264         heavy snow 1.793259e+10
```

```
crop_dmg_events[, c("EVTYPE", "crop_dmg")]
```

```
##                EVTYPE    crop_dmg
## 77            drought 13972566000
## 148             flood  5661968450
## 515       river flood  5029459000
## 377         ice storm  5022113500
## 203              hail  3025974480
## 352         hurricane  2741910000
## 361 hurricane typhoon  2607872800
```

```
## 132       flash flood  1421317100
## 118      extreme cold  1312973000
## 179      frost freeze  1094186000
```

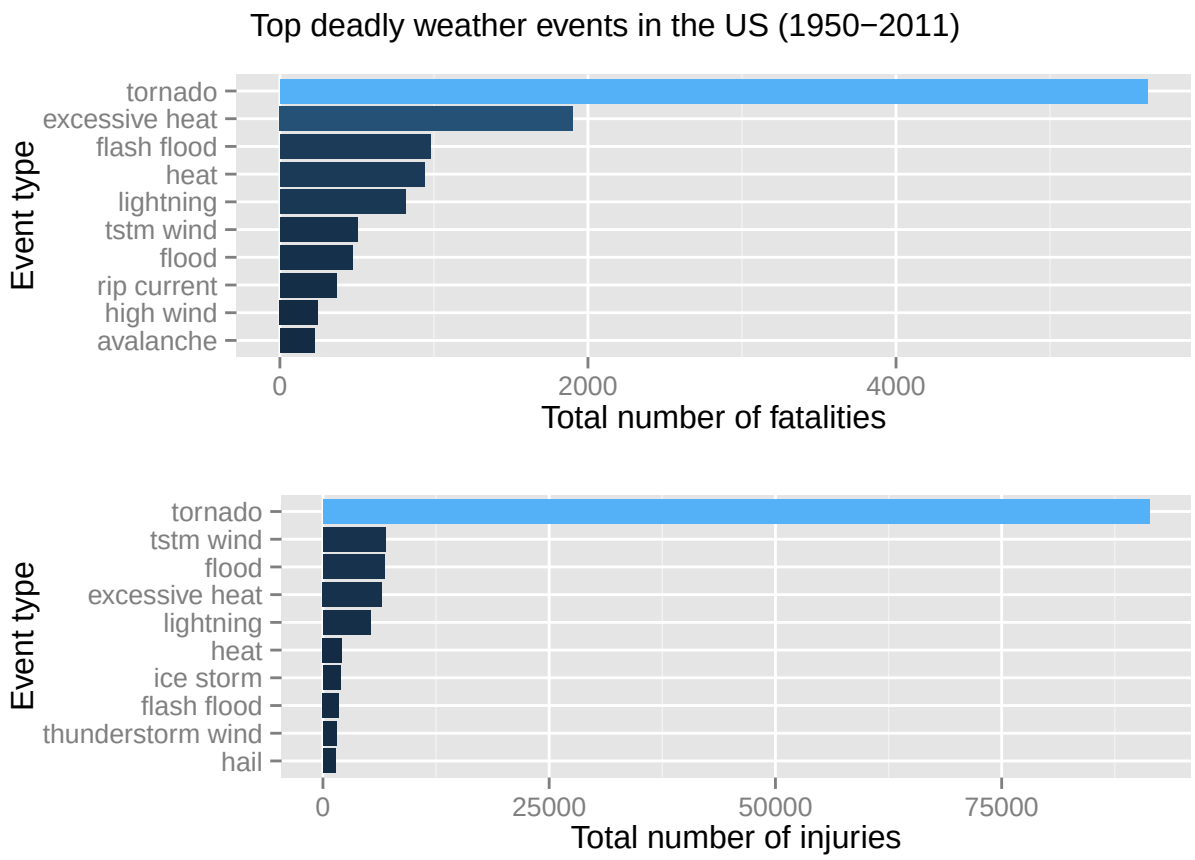
Results

Health impact of weather events

```
barplot1 <- ggplot(data=fatal,
             aes(x=reorder(EVTYPE, fatalities), y=fatalities, fill=fatalities)) +
    geom_bar(stat="identity") +
    coord_flip() +
    ylab("Total number of fatalities") +
    xlab("Event type") +
    theme(legend.position="none")

barplot2 <- ggplot(data=injury,
             aes(x=reorder(EVTYPE, injuries), y=injuries, fill=injuries)) +
    geom_bar(stat="identity") +
    coord_flip() +
    ylab("Total number of injuries") +
    xlab("Event type") +
    theme(legend.position="none")

grid.arrange(barplot1, barplot2, main="Top deadly weather events in the US (1950-2011)")
```

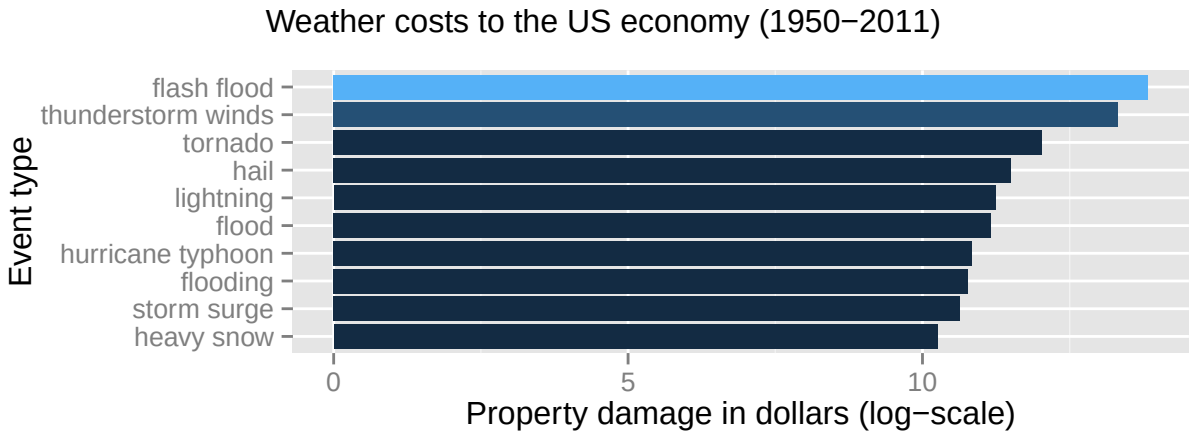


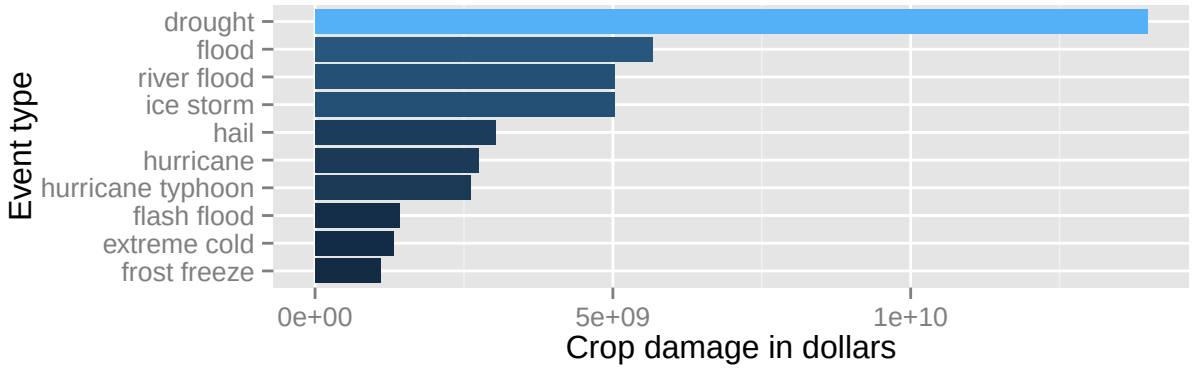
Tornadoes cause most number of deaths and injuries among all event types. There are more than 5,000 deaths and more than 10,000 injuries in the last 60 years in US, due to tornadoes. The other event types that are most dangerous with respect to population health are excessive heat and flash floods.

```
barplot1 <- ggplot(data=prop_dmg_events,
              aes(x=reorder(EVTYPE, prop_dmg), y=log10(prop_dmg), fill=prop_dmg )) +
    geom_bar(stat="identity") +
    coord_flip() +
    xlab("Event type") +
    ylab("Property damage in dollars (log-scale)") +
    theme(legend.position="none")

barplot2 <- ggplot(data=crop_dmg_events,
              aes(x=reorder(EVTYPE, crop_dmg), y=crop_dmg, fill=crop_dmg)) +
    geom_bar(stat="identity") +
    coord_flip() +
    xlab("Event type") +
    ylab("Crop damage in dollars") +
    theme(legend.position="none")

grid.arrange(barplot1, barplot2, main="Weather costs to the US economy (1950-2011)")
```





Property damages are given in logarithmic scale due to large range of values. The data shows that flash floods and thunderstorm winds cost the largest property damages among weather-related natural diseasters. Note that, due to untidy nature of the available data, type flood and flash flood are separate values and should be merged for more accurate data-driven conclusions.

The most severe weather event in terms of crop damage is the drought. In the last half century, the drought has caused more than 10 billion dollars damage. Other severe crop-damage-causing event types are floods and hails.