Analysis of natural events impact on Population and Economy

Synopsis

Using the National Weather Service Storm Data, we will verify what type of natural event has the most impact on population as injuries and deaths, and on economy as property damage and crop damage.

Data Processing

Download and import data

This code download the data from the original place and insert in the environment on the 'data' variable.

```
setwd("~/Google Drive/deep/coursera_data_science/Course 05 Reproducible Research/5-4-1-project")
download.file('https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2', 'data.csv.bz2
library(readr)
data <- read_csv("data.csv.bz2")
```

```
## 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.
```

Results

We'll show that Tornados and Floods are the most dangerous events, one for the population directly, through injuries, the other for damage to phisical properties and crops.

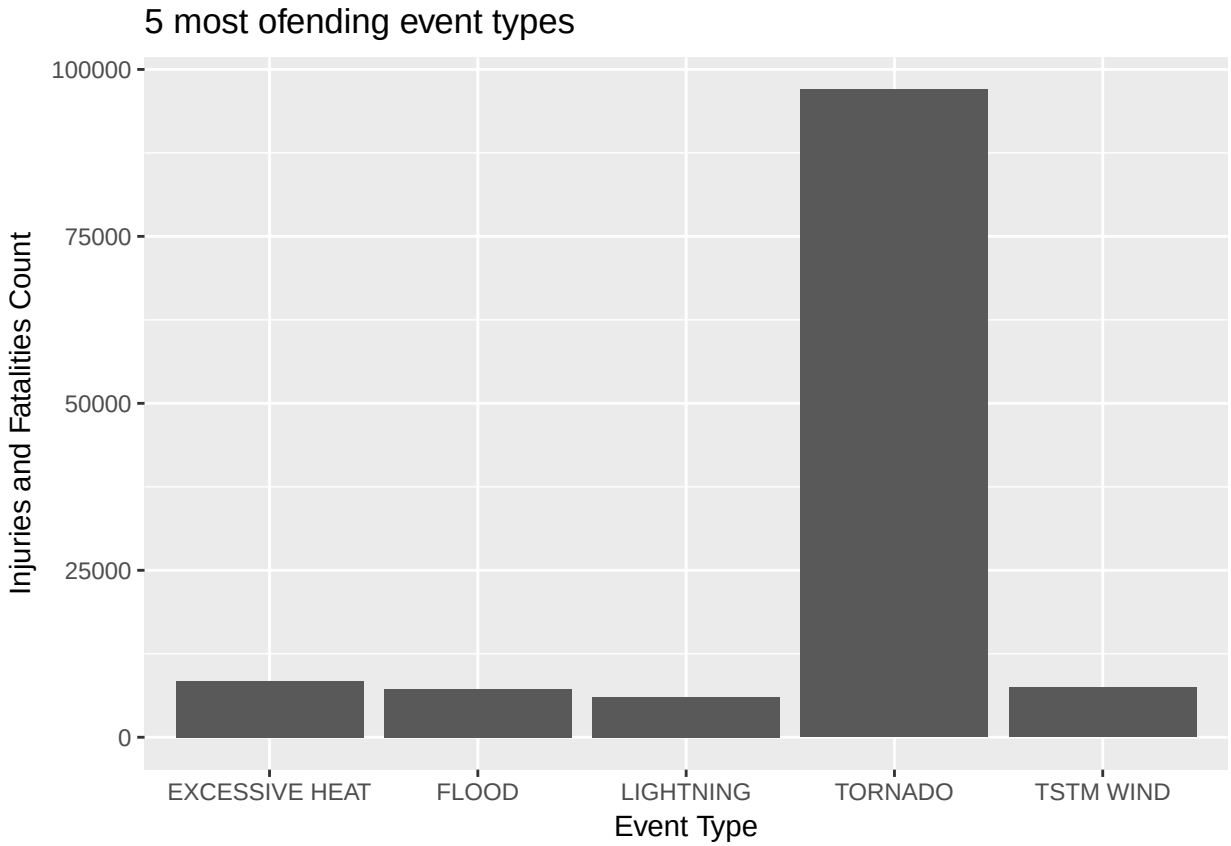
Injuries and Fatalities

First we just sum the two columns of injuries and fatalities. They'll have the same length of Event Type, which we will aggregate summing each one. It will be a data frame of two columns: Event Type and Health Count. Also, order in decreasing order (biggest injurers up).

```
total <- c(data$INJURIES) + c(data$FATALITIES)
human <- aggregate(total, by=list(data$EVTYPE), sum)
names(human) <- c("type", "count")
human <- human[order(human$count,decreasing = TRUE),]
```

Plot the aggressors. Tornados are by far the most dangerous event.

```
library(ggplot2)
ggplot(human[1:5,], aes(type, count)) + geom_col() +
  labs(title="5 most ofending event types") +
  labs(x='Event Type', y="Injuries and Fatalities Count")
```



Percentage of injuries and fatalities from tornados

```
human[1,2]/sum(human$count)
```

```
## [1] 0.6229661
```

Damage to properties and crops

```
# prepare a column with right numerator for property damage.
prop <- data$PROPDMG
size <- data$PROPDMGEXP
```

```
size[size=="K"] <- 10^3
size[size=="M"] <- 10^6
size[size=="B"] <- 10^9
size = as.numeric(size)
```

```
## Warning: NAs introduced by coercion
```

```
size[is.na(size)] <- 1
PROP <- prop * size

crop <- data$CROPDMG
size <- data$CROPDMGEXP
size[size=="K"] <- 10^3
size[size=="M"] <- 10^6
size[size=="B"] <- 10^9
size = as.numeric(size)
```

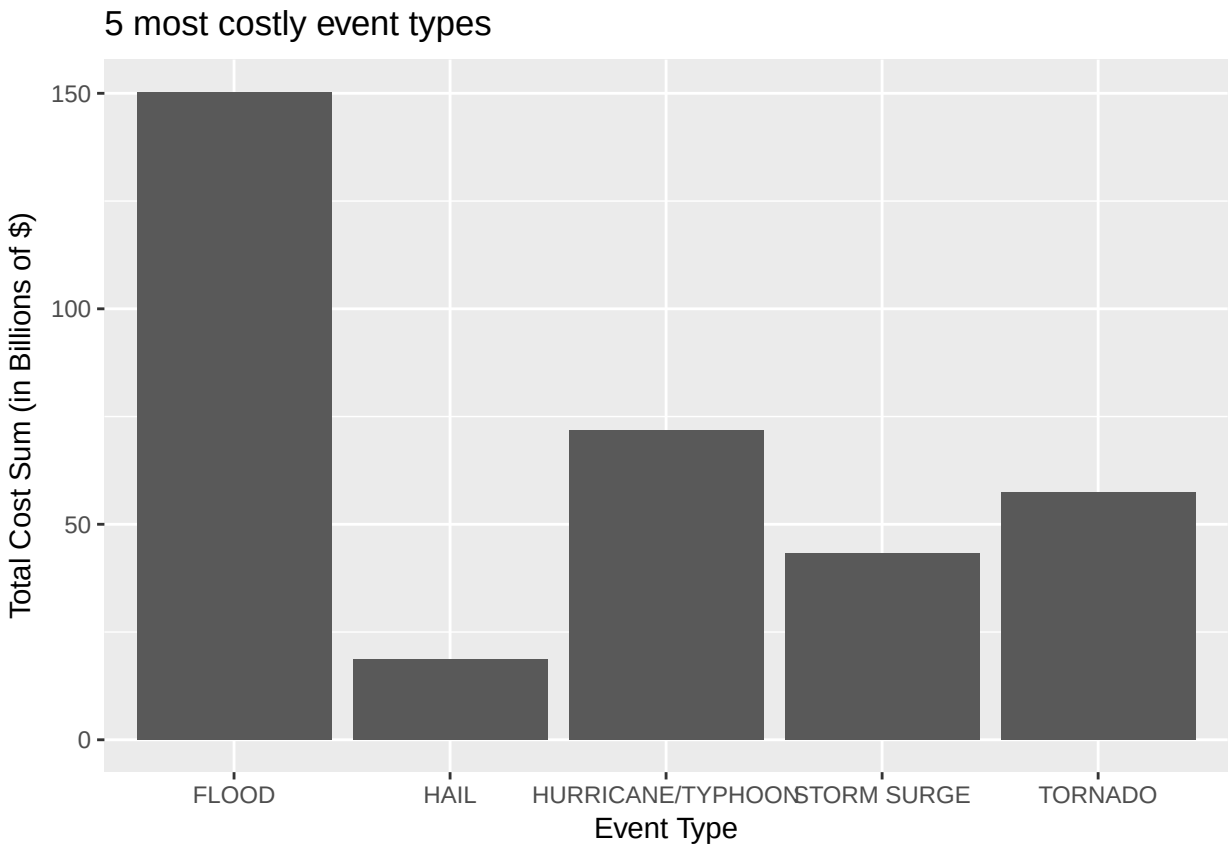
```
## Warning: NAs introduced by coercion
```

```
size[is.na(size)] <- 1
CROP <- crop * size

COST <- (CROP + PROP)/10^9

cost <- aggregate(COST, by=list(data$EVTYPE), sum)
names(cost) <- c("type", "sum")
cost <- cost[order(cost$sum, decreasing = TRUE),]

ggplot(cost[1:5,], aes(type, sum)) + geom_col() +
  labs(title="5 most costly event types") +
  labs(x='Event Type', y="Total Cost Sum (in Billions of $)")
```



Percentage of Flood to the total.

```
cost[1,2]/sum(cost$sum)
```

```
## [1] 0.31555
```