

Weather events that are most harmful to public health and have greatest economic consequences

Synopsis

Storms and other severe weather events have huge impact on public health and create economic problems for municipalities and their inhabitants. Some of the severe events can cause injuries, property damage and even lead to death. This analysis addresses which types of events are most harmful with respect to population health and which have the greatest economic consequences.

Data Processing

I am exploring the NOAA storm database which tracks characteristics of major storms and weather events in the United States. This dataset comes from the Internet.

Download file from the Internet:

```
filename <- 'C:/Users/Deep/Documents/StormData.csv.bz2'
if (!file.exists(filename)) {
  download.file('https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2', filename)
}
storm_data <- read.csv(filename)
# Ensure we got the data downloaded, decompressed and loaded correctly
# by checking filesize and dataset dimensions
stopifnot(file.size(filename) == 49177144)
stopifnot(dim(storm_data) == c(902297,37))
```

Read file in table format

```
StormData <- read.csv(bzfile("C:/Users/Deep/Documents/StormData.csv.bz2"), sep = ",", header=TRUE)
```

Property damage estimates were entered as actual dollar amounts (the variable PROPDMG). But they were rounded to three significant digits, followed by an alphabetical character signifying the magnitude of the number, i.e., 1.55B for \$1,550,000,000. Alphabetical characters used to signify magnitude include “K” for thousands, “M” for millions, and “B” for billions. So I created a new variable PROPDMGEXP2 and assigned conditionally “K” = 1000, “M” = 1000000, “B” = 1000000000, in other cases 1. These variables are multiplied in the next step.

```
table(StormData$PROPDMGEXP)
```

```
##
##      -      ?      +      0      1      2      3      4      5
## 465934  1      8      5    216    25     13     4      4     28
##      6      7      8      B      h      H      K      m      M
##      4      5      1     40      1      6 424665    7   11330
```

```
StormData$PROPDMGEXP2 <- 1
StormData$PROPDMGEXP2[which(StormData$PROPDMGEXP == "K")] <- 1000
StormData$PROPDMGEXP2[which(StormData$PROPDMGEXP == "M" | StormData$PROPDMGEXP == "m")] <- 1000000
StormData$PROPDMGEXP2[which(StormData$PROPDMGEXP == "B")] <- 1000000000
```

```
table(StormData$PROPDMGEXP2)
```

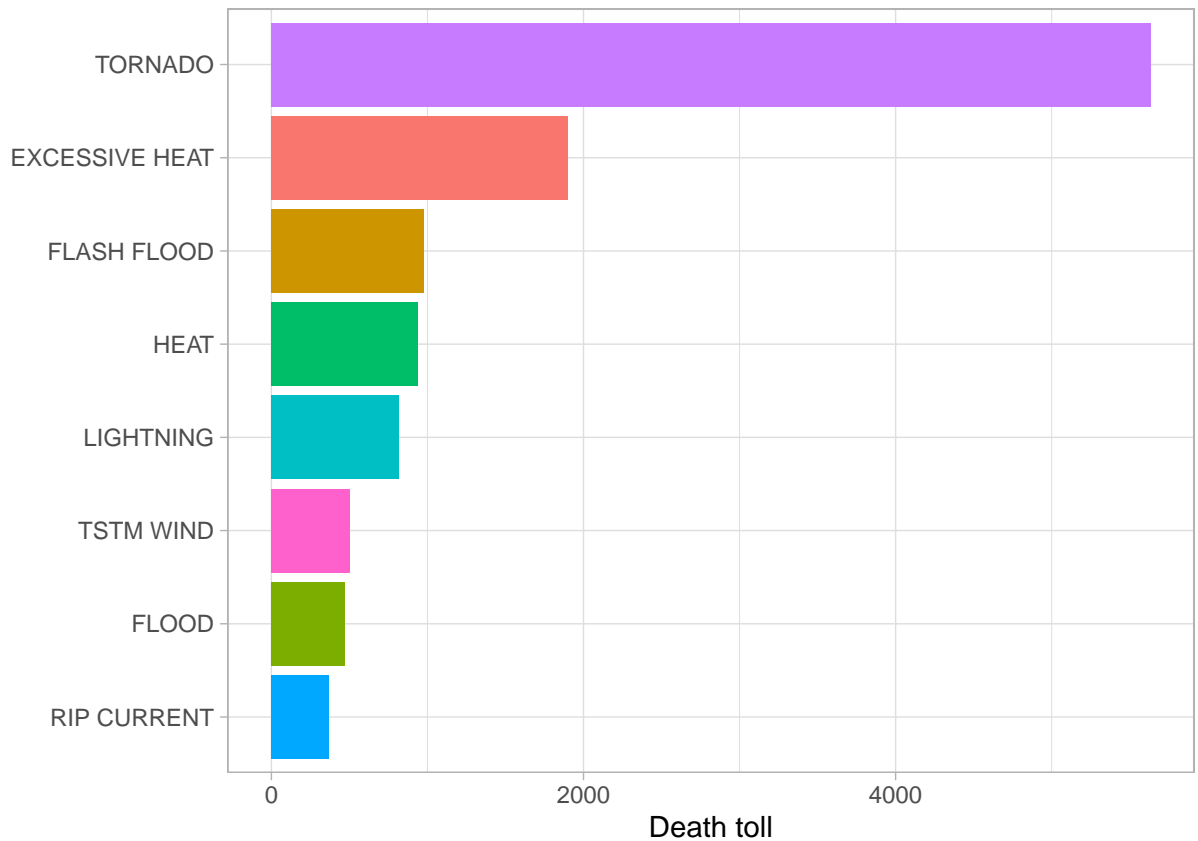
```
##
##      1      1000    1e+06    1e+09
## 466255 424665    11337         40
```

Which types of events are most harmful to population health?

Fatalities and injuries have the most impact on public health, so I will present what types of severe weather are the most dangerous.

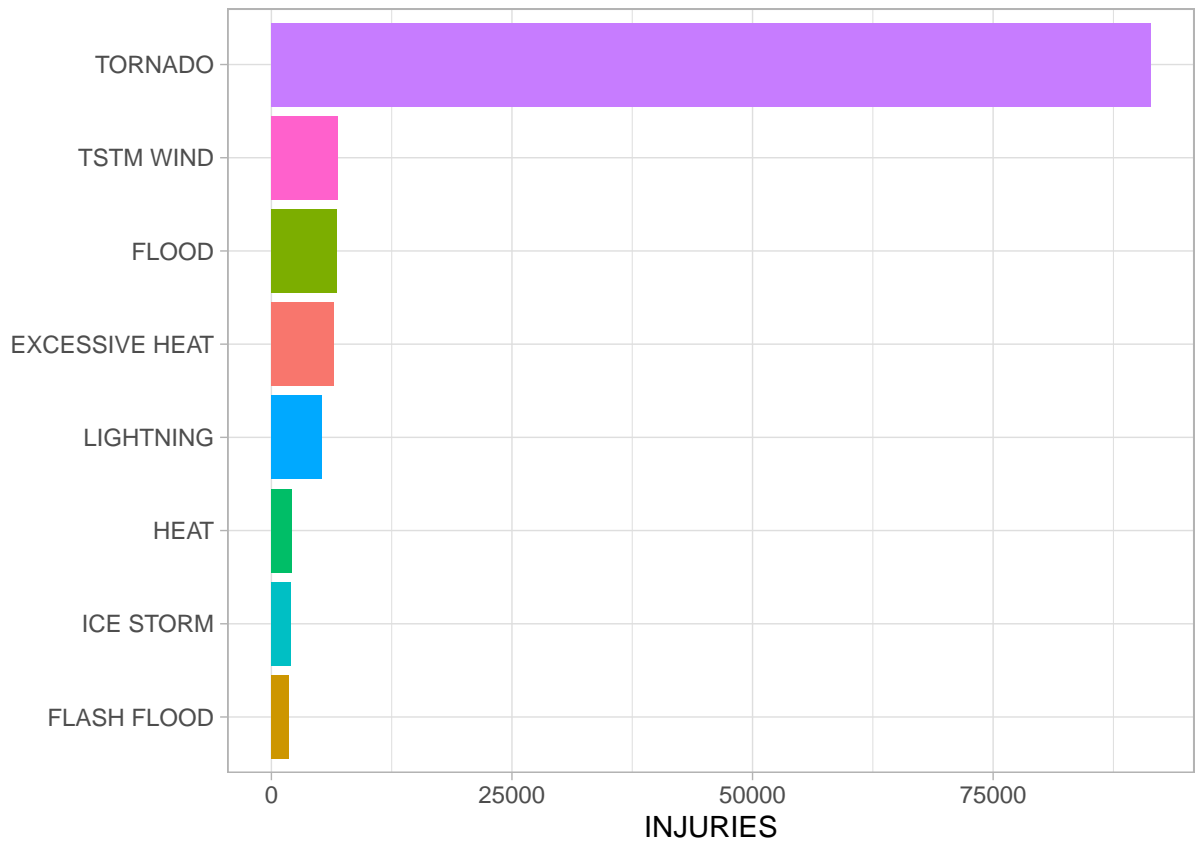
The first plot presents a Death toll by Event type

```
StormData %>%
  select(FATALITIES, EVTYPE) %>%
  group_by(EVTYPE) %>%
  summarise(SumFATALITIES = sum(FATALITIES)) %>%
  top_n(n = 8, wt = SumFATALITIES) %>%
  ggplot(aes(y = SumFATALITIES, x = reorder(x = EVTYPE, X = SumFATALITIES), fill=EVTYPE))+
  geom_bar(stat = "identity", show.legend = FALSE) +
  #geom_text(aes(label=SumFATALITIES), size = 4, hjust = 0.5, vjust = -0.1) +
  xlab(label = "") +
  ylab(label = "Death toll") +
  coord_flip() +
  theme_light()
```



The second plot presents Injuries by Event type

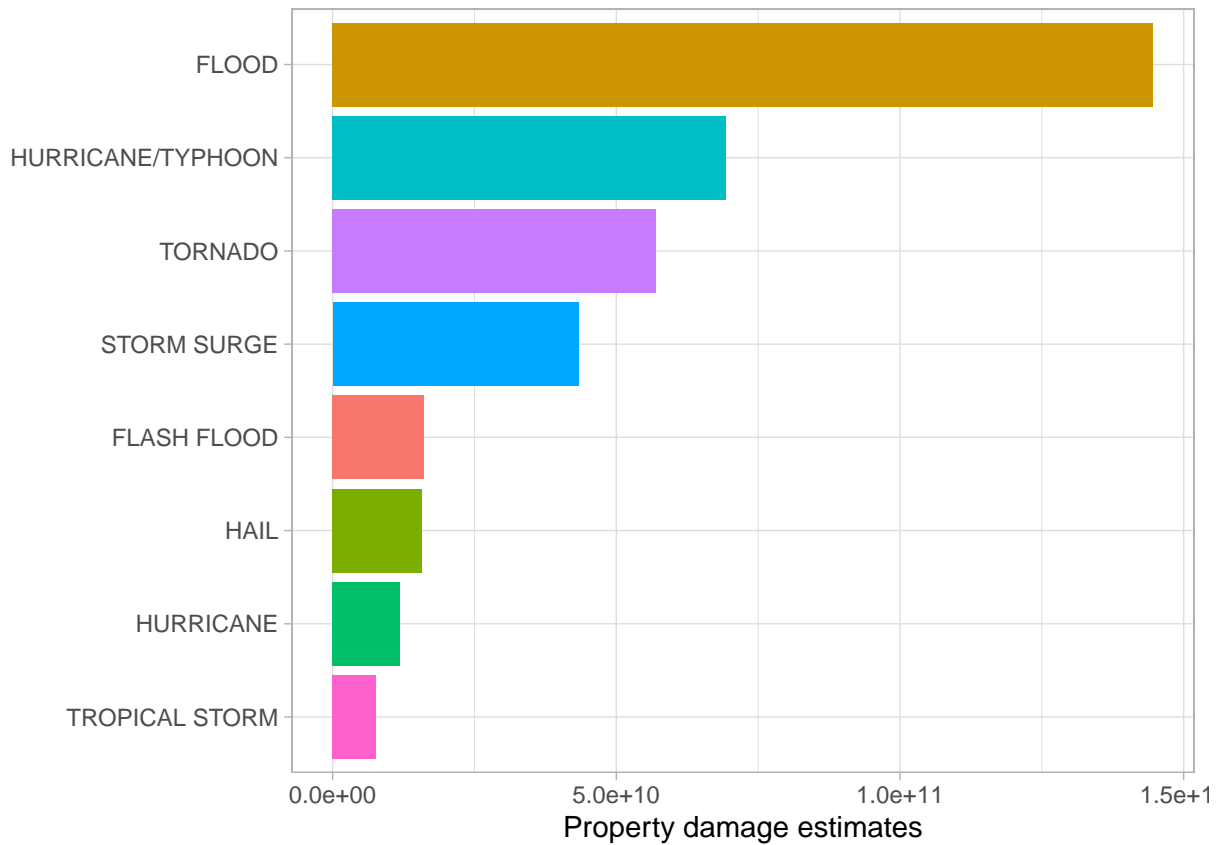
```
StormData %>%
  select(INJURIES, EVTYPE) %>%
  group_by(EVTYPE) %>%
  summarise(SumINJURIES = sum(INJURIES)) %>%
  top_n(n = 8, wt = SumINJURIES) %>%
  ggplot(aes(y = SumINJURIES, x = reorder(x = EVTYPE, X = SumINJURIES), fill=EVTYPE)) +
  geom_bar(stat = "identity", show.legend = FALSE) +
  #geom_text(aes(label=SumINJURIES), size = 4, hjust = 0.5, vjust = -0.1) +
  xlab(label = "") +
  ylab(label = "INJURIES") +
  coord_flip() +
  theme_light()
```



Which types of events have the greatest economic consequences?

This plot shows Property damage estimates by Event type

```
StormData %>%
  select(PropDMG, PropDMGEXP2, EVTYPE) %>%
  group_by(EVTYPE) %>%
  mutate(SumPropDMGEXP = (PropDMG * PropDMGEXP2)) %>%
  summarise(SumPropDMGEXP2 = sum(SumPropDMGEXP)) %>%
  top_n(n = 8, wt = SumPropDMGEXP2) %>%
  ggplot(aes(y = SumPropDMGEXP2, x = reorder(x = EVTYPE, X = SumPropDMGEXP2), fill=EVTYPE)) +
  geom_bar(stat = "identity", show.legend = FALSE) +
  #geom_text(aes(label=SumFATALITIES), size = 4, hjust = 0.5, vjust = -0.1) +
  xlab(label = "") +
  ylab(label = "Property damage estimates") +
  coord_flip() +
  theme_light()
```



Conclusion

As you can see above flood has the greatest economic consequences. Tornado is the most harmful to population health because caused the most death tolls and injuries.