

Economic and Medical Impact of Extreme Weather

S. Smolenski

Synopsis

We have considered the economic impact, as indicated by property damage and crop damage, along with the harm to population, as indicated by fatalities and injuries, caused by different types of weather events. The data come from the NOAA in 2007. Throughout this document, I will use the word ‘cost’ in a general way to mean both monetary cost and casualty cost.

Data Processing

I first read in the zip files from the course website and unzip:

```
datfile<-"C:/Users/Sarah/Downloads/StormData.csv"
if (!file.exists(datfile)){
  url<-"https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2"
  download.file(url, "StormData.zip")
  unzip("StormData.zip")
}
```

For the processing, I will use the following libraries:

```
library(dplyr)
library(ggplot2)
library(tidyr)
```

Then I read in the csv file, selecting only the data of interest to this analysis (ie, the data on fatalities, injuries, and damage costs). I group the data by event type, so that I can take the mean of each cost column, and keep only the event types that have non-zero values for at least one cost column.

```
read.csv(datfile) %>%
  select(EVTYPE,FATALITIES,INJURIES,PROPDMG,CROPDMG) %>%
  group_by(EVTYPE) -> data

summarize( data,
  FATALITIES=mean(FATALITIES),
  INJURIES=mean(INJURIES),
  PROPDMG=mean(PROPDMG),
  CROPDMG=mean(CROPDMG)) %>%
  filter(FATALITIES!=0 & INJURIES!=0 & PROPDMG!= 0 & CROPDMG!=0) -> means
```

I then combine the injuries and fatalities columns, creating a ‘health cost’ table:

```
gather(means,"Type","Number",2:3) %>%
  select(EVTYPE,Type,Number) %>%
  arrange(desc(Number)) -> health
```

And take only the top ten events on this table:

```
htypes<-unique(health$EVTYPE)[1:10]
```

I then do the same thing for the economic costs.

```
gather(means, "Type", "Number", 4:5) %>%
  select(EVTYPE, Type, Number) %>%
  arrange(desc(Number)) -> property

ptypes<-unique(property$EVTYPE)[1:10]
```

At this point, the data is in a convenient form to visualize.

Results

Health Impact

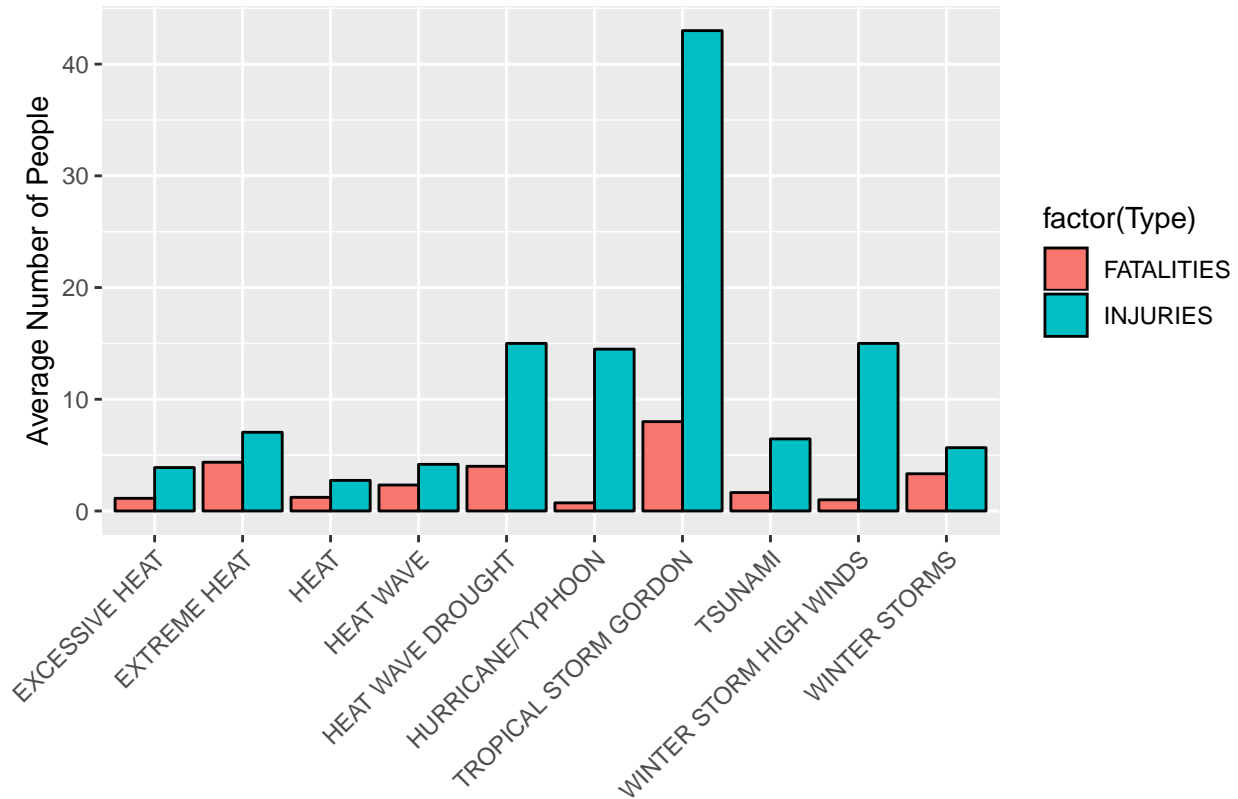
The ten weather events that have the highest impact on population health are:

```
htypes
```

```
## [1] TROPICAL STORM GORDON    HEAT WAVE DROUGHT
## [3] WINTER STORM HIGH WINDS HURRICANE/TYPHOON
## [5] EXTREME HEAT            TSUNAMI
## [7] WINTER STORMS           HEAT WAVE
## [9] EXCESSIVE HEAT          HEAT
## 985 Levels:    HIGH SURF ADVISORY COASTAL FLOOD ... WND
```

Visualizing these events, we have:

```
ggplot(subset(health, EVTYPE %in% htypes), aes(x=EVTYPE, y=Number, fill=factor(Type))) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  geom_bar(stat="identity", position="dodge", color="black") +
  scale_color_discrete("Type") +
  xlab("") + ylab("Average Number of People")
```



And we can see clearly that Tropical Storm Gordon caused both the most fatalities and the most injuries.

Economic Impact

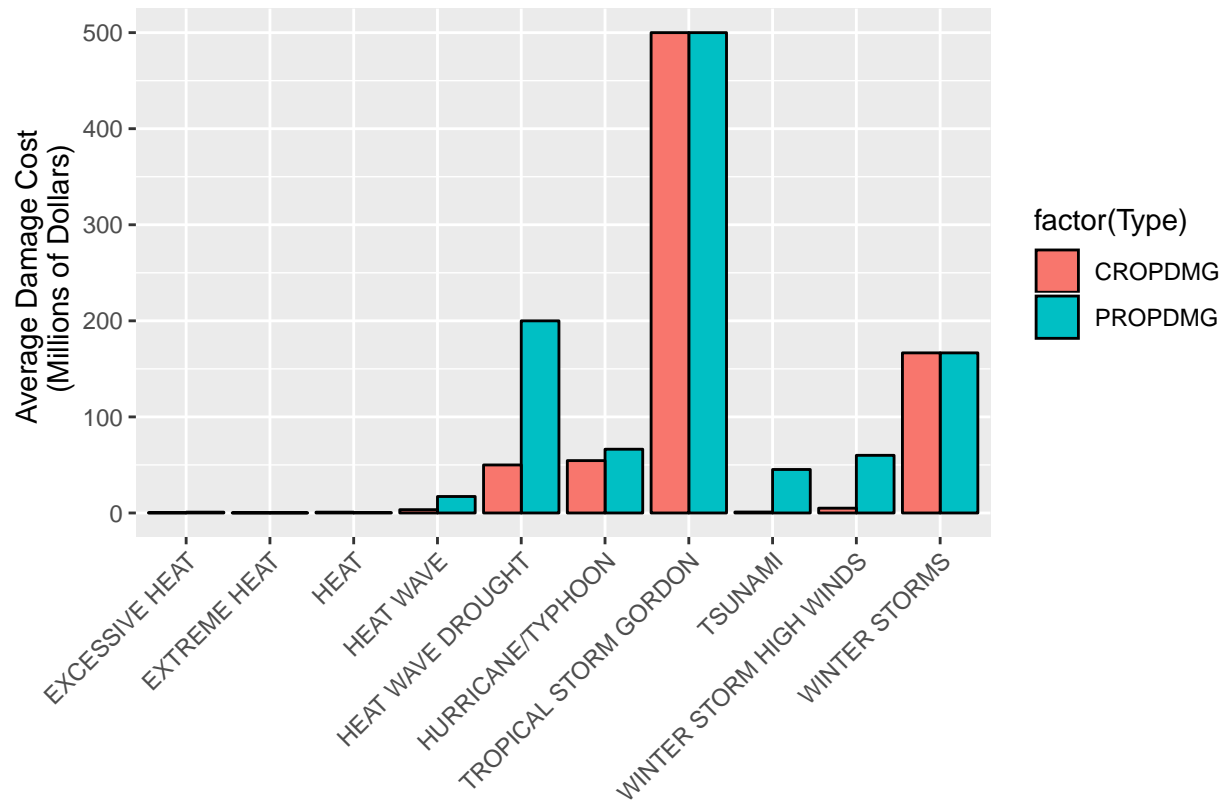
The ten weather events that have the highest impact on population health are:

ptypes

```
## [1] TROPICAL STORM GORDON   HEAT WAVE DROUGHT
## [3] WINTER STORMS           HURRICANE OPAL
## [5] HURRICANE               RIVER FLOOD
## [7] STORM SURGE             TROPICAL STORM
## [9] HURRICANE/TYPHOON      WINTER STORM HIGH WINDS
## 985 Levels:   HIGH SURF ADVISORY COASTAL FLOOD ... WND
```

Visualizing these events, we have:

```
ggplot(subset(property, EVTYPE %in% htypes), aes(x=EVTYPE, y=Number, fill=factor(Type))) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  geom_bar(stat="identity", position="dodge", color="black") +
  scale_color_discrete("Type") +
  xlab("") + ylab("Average Damage Cost \n (Millions of Dollars)")
```



And, once again, we can see clearly that Tropical Storm Gordon caused both the most property damage and the most crop damage.