Reproducible Research Week4 Project: Analyses of Storms on Population Health and Economy

Akira Sano

1/7/2020

Synopsis

We explore the NOAA Storm Database and answer some basic questions about severe weather events. We concentrate on analysing the storm events that affect both population health and economy. We present our results to illustrate which storm events in the past have cause the most damage to human population and health, as well as to our economy in properties and farm crops.

Date Processing

We use R programming for data processing and analysis, then publish in pdf.

We first import the data, transform the date variable to R's Date class, and find the oldest and most recent records.

```
# Download data
library(data.table)
fileUrl <- "https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2"
download.file(fileUrl, destfile = "stormdata.csv.bz2")
data0 <- fread("stormdata.csv.bz2")
data0$BGN_DATE <- as.Date(data0$BGN_DATE, "%m/%d/%Y %H:%M:%S")
data0[which.min(data0$BGN_DATE), BGN_DATE]
```

```
## [1] "1950-01-03"
```

```
data0[which.max(data0$BGN_DATE), BGN_DATE]
```

```
## [1] "2011-11-30"
```

We find that the avialble data spans from 1950 to 2011.

Results

We will analyze the data to answer the following two questions.

1. Which types of events are most harmful to population health?
2. Which types of events have the greatest economic consequences?

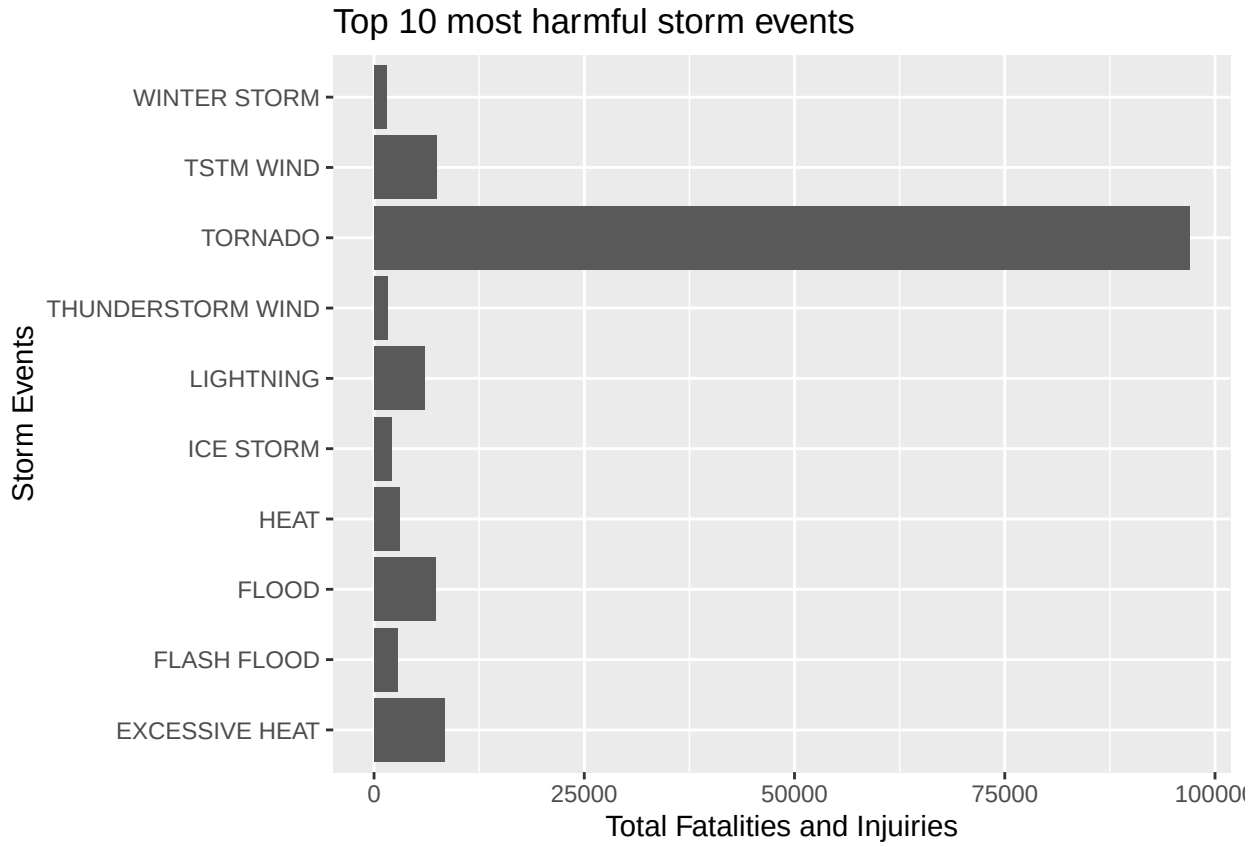
Which types of events are most harmful to population health?

We interpret the population health to be the sum of total fatalities and injuries, which the data set has, and decided that the most harmful storm event to be the one with the highest number of fatalities and injuries combined, caused by the event.

```
data1 <- data0[, .(sum_fatal_inj = sum(FATALITIES, INJURIES)), by=EVTYPE]
data1 <- data1[order(-sum_fatal_inj), ]
```

We plot the top 10 events in the sum of fatalities and injuries.

```
library(ggplot2)
ggplot(data1[1:10,], aes(x=EVTYPE, y=sum_fatal_inj)) + geom_col() + coord_flip() +
  xlab("Storm Events") + ylab("Total Fatalities and Injuiries") +
  ggtitle("Top 10 most harmful storm events")
```



Numerically, we have

```
data1[which.max(data1[, sum_fatal_inj]), ]
```

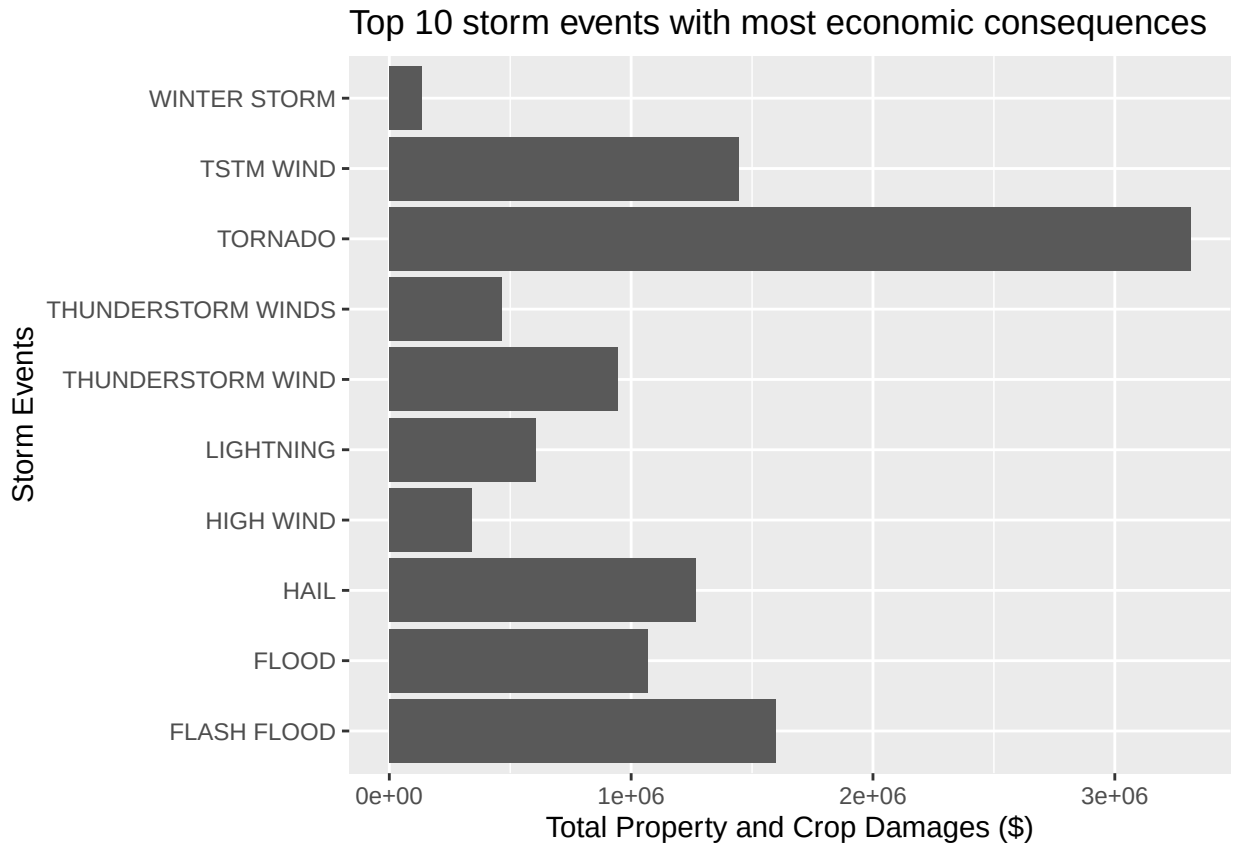
```
##     EVTYPE sum_fatal_inj
## 1: TORNADO         96979
```

We conclude that with the fatalities and injuries combined total of 96,979 since 1950, tornados have been the most harmful storm event to population health.

Which types of events have the greatest economic consequences?

We define the economic consequence as the dollar value given by the sum of the property damage and crop damage, both given in the data set. We plot the 10 storm events with higest economic consequences.

```
data2 <- data0[, .(sum_prop_crop_dmgs = sum(PROPDMG, CROPDMG)), by=EVTYPE]
data2 <- data2[order(-sum_prop_crop_dmgs)]
ggplot(data2[1:10,], aes(x=EVTYPE, y=sum_prop_crop_dmgs)) + geom_col() + coord_flip() +
  xlab("Storm Events") + ylab("Total Property and Crop Damages ($)") +
  ggtitle("Top 10 storm events with most economic consequences")
```



Numerically, we have

```
data2[which.max(data2[, sum_prop_crop_dmgs]), ]
```

```
##     EVTYPE sum_prop_crop_dmgs
## 1: TORNADO            3312277
```

We conclude that with the property and crop damage combined total of $3,312,277 since 1950, tornados have again caused the most economic consequences.

We summarize that tornados have been the storm event that have cuased the most damage, both to population and to economy.