Storm Data Analysis

Kazuyuki Ishihara

2022-09-05

Synopsis

This document is a result of the analysis of "Storm Data" that is published from the U.S. National Oceanic and Atmospheric Administration's (NOAA) storm database. The results suggested that "TORNADO" is the most harmful event both on population health and economic damages.

Processing Data

Qustion 1: Across the United States, which types of events (as indicated in the EVTYPE variable) are most harmful with respect to population health?

I describe the steps of processing data as followings.
1. read the file.
2. extract necessary columns for subsequent data analysis.
3. calculate the number of total health influences(variables "FATALITIES"+"INJURIES") by each event (EVTYPE).

```
Storm <- read.csv("repdata_data_StormData.csv")
library(dplyr)
StormEv <- select(Storm, EVTYPE, FATALITIES, INJURIES, PROPDMG)
EVHealth <- aggregate(FATALITIES+INJURIES~EVTYPE, data = StormEv, sum)
```

Qustion 2: Across the United States, which types of events have the greatest economic consequences?

4. calculate the number of total economic damages(variable "PROPDMG") by each event (EVTYPE).

```
EVEcon <- aggregate(PROPDMG~EVTYPE, data = StormEv, sum)
```

Results

Qustion 1: Across the United States, which types of events (as indicated in the EVTYPE variable) are most harmful with respect to population health?

```
head(EVHealth[order(EVHealth$`FATALITIES + INJURIES`, decreasing = T),])
```

```
##             EVTYPE FATALITIES + INJURIES
## 834        TORNADO                 96979
## 130 EXCESSIVE HEAT                  8428
```

```
## 856      TSTM WIND                   7461
## 170          FLOOD                   7259
## 464      LIGHTNING                   6046
## 275           HEAT                   3037
```

The table shows that "TORNADO" induced the maximum number of population health damages.

Qustion 2: Across the United States, which types of events have the greatest economic consequences?

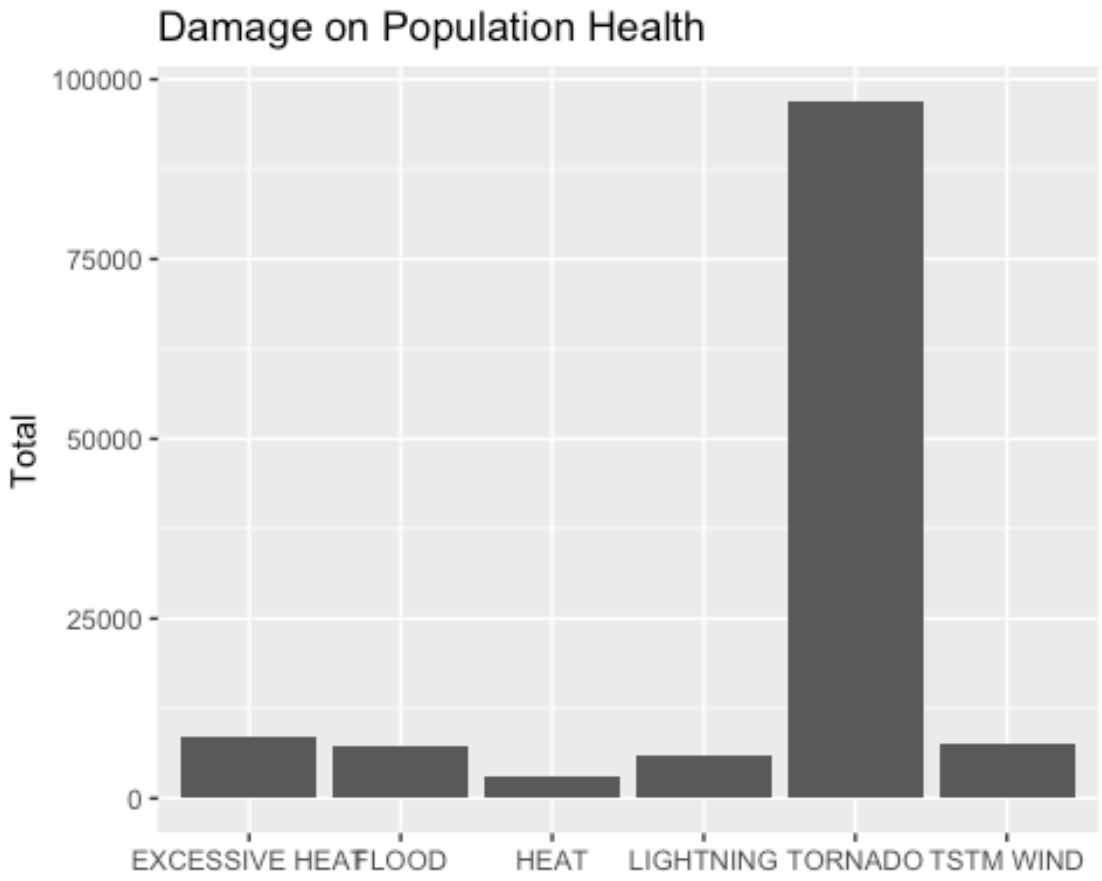
```
head(EVEcon[order(EVEcon$PROPDMG, decreasing = T),])
```

```
##                EVTYPE   PROPDMG
## 834           TORNADO 3212258.2
## 153       FLASH FLOOD 1420124.6
## 856         TSTM WIND 1335965.6
## 170             FLOOD  899938.5
## 760 THUNDERSTORM WIND  876844.2
## 244              HAIL  688693.4
```

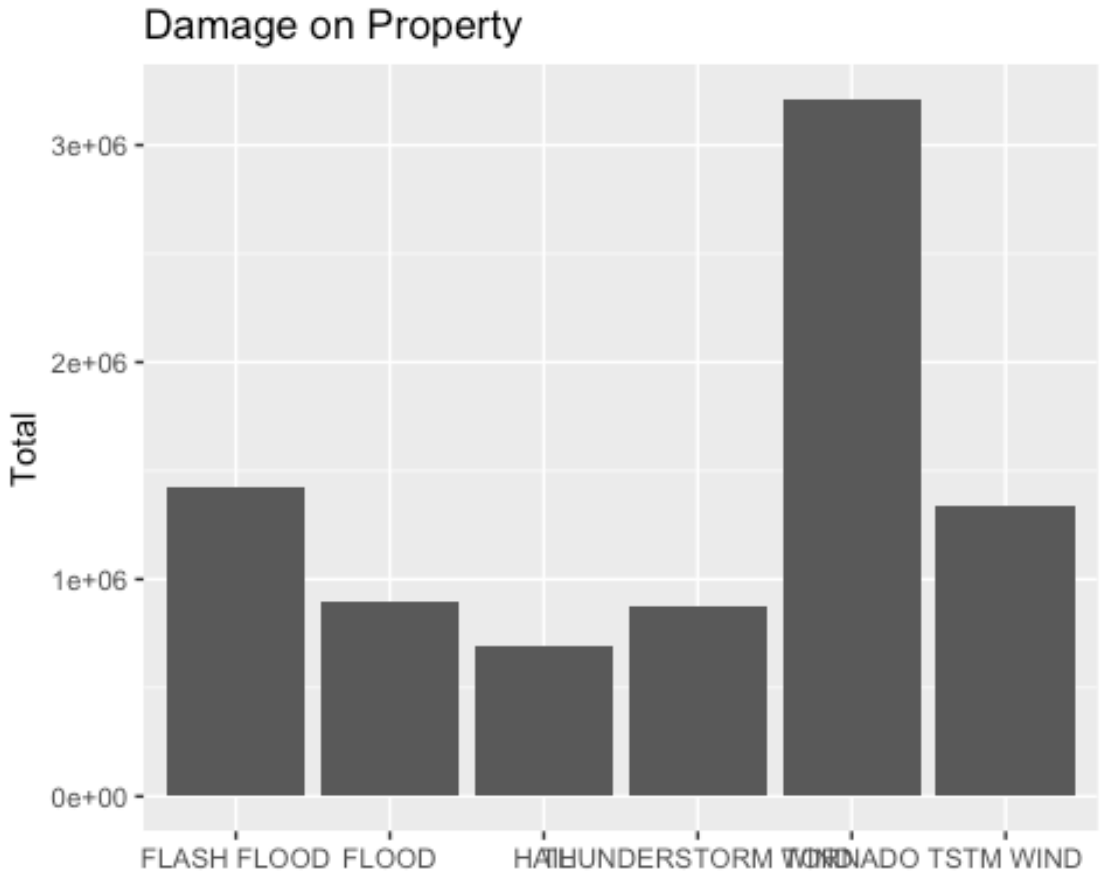
The table shows that "TORNADO" also induced the maximum number of property damages.

Data plot of top 6 events of each question.

```
Top6Event <- data.frame("Health" = head(EVHealth[order(EVHealth$`FATALITI
ES + INJURIES`, decreasing = T),]), "Economy" = head(EVEcon[order(EVEcon
$PROPDMG, decreasing = T),]))
library(ggplot2)
g1 <- ggplot(Top6Event, aes(Health.EVTYPE, Health.FATALITIES...INJURIES))
g1 + geom_bar(stat = "identity") + labs(x = NULL, y = "Total", title = "D
amage on Population Health")
```




```
g2 <- ggplot(Top6Event, aes(Economy.EVTYPE, Economy.PROPDMG))
g2 + geom_bar(stat = "identity") + labs(x = NULL, y = "Total", title = "D
amage on Property")
```


Damage on Property
Total
3e+06
2e+06
1e+06
0e+00
FLASH FLOOD
FLOOD
HAIL
THUNDERSTORM WIND
TORNADO
TSTM WIND