

Analysis on Severe Weather Impact to Population and Economic in the US

Fadhel Hariz Dzulfar
11/1/2020

Synopsis

This project attempts to determine the most harmful weather impact on public health and economy in the US by exploring the U.S. National Oceanic and Atmospheric Administration's (NOAA) storm database. This database tracks characteristics of major storms and weather events in the United States, including when and where they occur, as well as estimates of any fatalities, injuries, and property damage.

Data Processing

```
library(R.utils)
library(plyr)
library(ggplot2)
library(dplyr)
```

Download the data

```
if (!file.exists("storm_data.csv")) {
  file_url <- "https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2"
  file_name <- "storm_data.csv.bz2"
  download.file(file_url, file_name)

  bunzip2(file_name)
}
```

Read the data

```
storm_data <- read.csv('storm_data.csv')
dim(storm_data)
```

```
## [1] 902297    37
```

Convert the variable PROPDMG from the character to actual values ("K"=1000, "M"=1000000, "B"=1000000000) and assign them into a new variable PROPDMG2

```
table(storm_data$PROPDMGEXP)
```

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

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

```
table(storm_data$PROPDMGEXP2)
```

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

Results

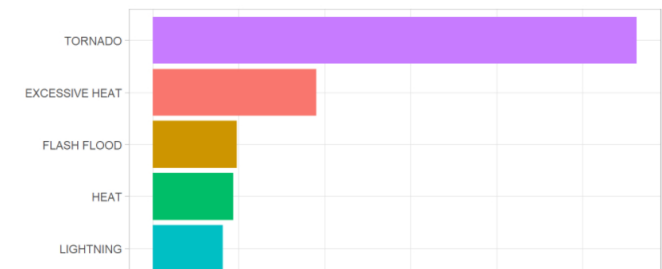
1. Across the United States, which types of events are most harmful with respect to population health?

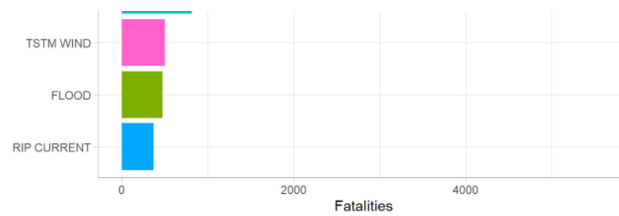
To answer this question, I will present what types of severe weather are the most dangerous based on fatalities and injuries variable.

The first plot will show the data based on fatalities

```
storm_data %>%
  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) +
  xlab(label = "") +
  ylab(label = "Fatalities") +
  coord_flip() +
  theme_light()

## `summarise()` ungrouping output (override with `.groups` argument)
```

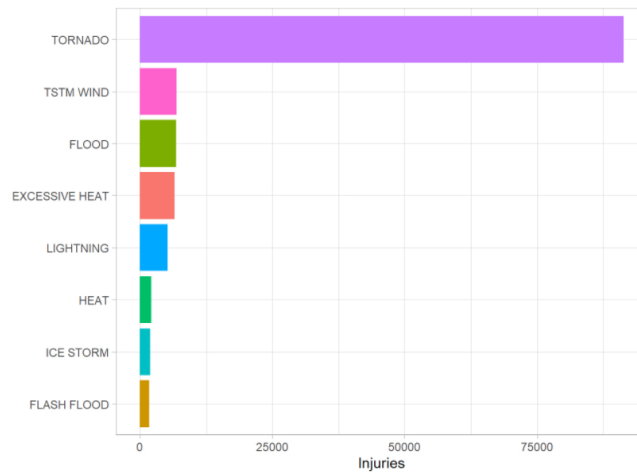




The second plot will show the data based on injuries

```
storm_data %>%
  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) +
  xlab(label = "") +
  ylab(label = "Injuries") +
  coord_flip() +
  theme_light()
```

`summarise()` ungrouping output (override with `.groups` argument)

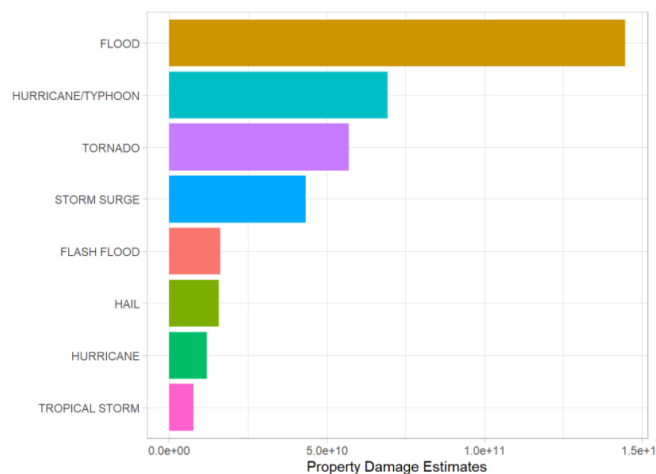


2. Across the United States, which types of events have the greatest economic consequences?

This plot will show the data based on property damage estimates.

```
storm_data %>%
  select(PROPDGM, PROPDMGEXP2, EVTYPE) %>%
  group_by(EVTYPE) %>%
  mutate(SumPROPDMGEXP = (PROPDGM * 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) +
  xlab(label = "") +
  ylab(label = "Property Damage Estimates") +
  coord_flip() +
  theme_light()
```

`summarise()` ungrouping output (override with `.groups` argument)



Conclusion

Based on the above plots, we can see that tornado events contributed the most to the number of fatalities and injuries. Therefore, tornado events are most harmful with respect to population health. On the other hand, the flood events contributed the most in terms of property damage

