Analyzing the Impact of Severe Weather Events

Kevin O'Brien

19 March 2023

Introduction

Storms and other severe weather events regularly cause public health disasters and economic hardships. These events often result in fatalities, injuries, and property damage.

This analysis is a brief exploration of the U.S. National Oceanic and Atmospheric Administration's (NOAA) storm database. This database tracks when and where events occur along with estimated impacts. The data can be found here: https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2.

The primary goal of this analysis is to answer the following two questions: 1. Across the United States, which types of events (as indicated in the EVTYPE variable) are most harmful with respect to population health? 2. Across the United States, which types of events have the greatest economic consequences?

Loading the data and relevant packages

The first step is to load the data. In this analysis I chose to use the tidyverse family of packages for data processing, including dplyr, readr and ggplot2. For plotting I used ggplot2.

Data Processing

Using fatalities and injuries the data was aggregated using the {tidyverse} family of R packages. Sorting by fatalities first then injuries and eliminating event without documented damage to human population.

Health Impact Aggregation

This segment of code selects the relevant data columns, aggregates fatalities and injuries then sorts and selects the top 10.

```
##Selects the columns which include type of disaster, fatality and injury counts.
storm_h <- storm %>% select(c("EVTYPE","FATALITIES" ,"INJURIES"))
```

```
##Using dplyr, I aggregate the sums for fatalities and injuries by event type and sorts by fatality cou

health_agg <- storm_h %>% group_by(EVTYPE) %>%
  summarize(Total_Fatalities = sum(FATALITIES),
            Total_Injuries = sum(INJURIES)) %>%
  arrange(desc(Total_Fatalities),desc(Total_Injuries))

##Considers only the top 10 most damaging events to human populations.
health_top10 <- health_agg %>%
  head(10)
```

Economic Impact Aggregation

This segment of code selects the relevant data columns, aggregates property damage and crop damage and then selects the top 25.

```
##Selects the columns with type of disaster, property damage and crop damage.
storm_e <- storm %>% select(c("EVTYPE", "PROPDMG"   , "CROPDMG"))

##Aggregates damage amounts by event type.
econ_agg <- storm_e %>% group_by(EVTYPE) %>%
  summarize(Total_Prop_Damage = sum(PROPDMG),
            Total_Crop_Damage  = sum(CROPDMG)) %>%
  arrange(desc(Total_Prop_Damage),desc(Total_Crop_Damage))

##Selects 25 event types with highest damage amounts
econ_top25 <- econ_agg[1:25,]
```

Health Impact Reformatting

The primary objective of this code is to create useful factors for graphing. The first portion deals with injury/fatality factorization. The 2nd segment assigns an order to the 10 most dangerous events.

```
health_top10
```

```
## # A tibble: 10 x 3
##    EVTYPE         Total_Fatalities Total_Injuries
##    <chr>                     <dbl>          <dbl>
##  1 TORNADO                    5633          91346
##  2 EXCESSIVE HEAT             1903           6525
##  3 FLASH FLOOD                 978           1777
##  4 HEAT                        937           2100
##  5 LIGHTNING                   816           5230
##  6 TSTM WIND                   504           6957
##  7 FLOOD                       470           6789
##  8 RIP CURRENT                 368            232
##  9 HIGH WIND                   248           1137
## 10 AVALANCHE                   224            170
```

```
##Reformats the data, which allows use of the injury/fatality distinction as a factor rather than a colu


IncidentRank <- health_top10 %>% pull(EVTYPE)
health_top10p <- health_top10 %>% pivot_longer(Total_Fatalities:Total_Injuries,names_to = "Type",values_
health_top10p <- health_top10p %>% mutate(EVTYPE = factor(EVTYPE,levels=IncidentRank))
```

```
head(health_top10p) %>% kable()
```

EVTYPE	Type	Value
TORNADO	Total_Fatalities	5633
TORNADO	Total_Injuries	91346

EVTYPE	Type	Value
EXCESSIVE HEAT	Total_Fatalities	1903
EXCESSIVE HEAT	Total_Injuries	6525
FLASH FLOOD	Total_Fatalities	978
FLASH FLOOD	Total_Injuries	1777

Economic Impact Data Reformatting

The primary objective of this code is to create useful factors for graphing. The first portion deals with property/crop damage factorization. The 2nd segment assigns an order to the 25 most dangerous events. The Damage values will be transformed into a variable denominated in millions of dollars, for the sake of legibility.

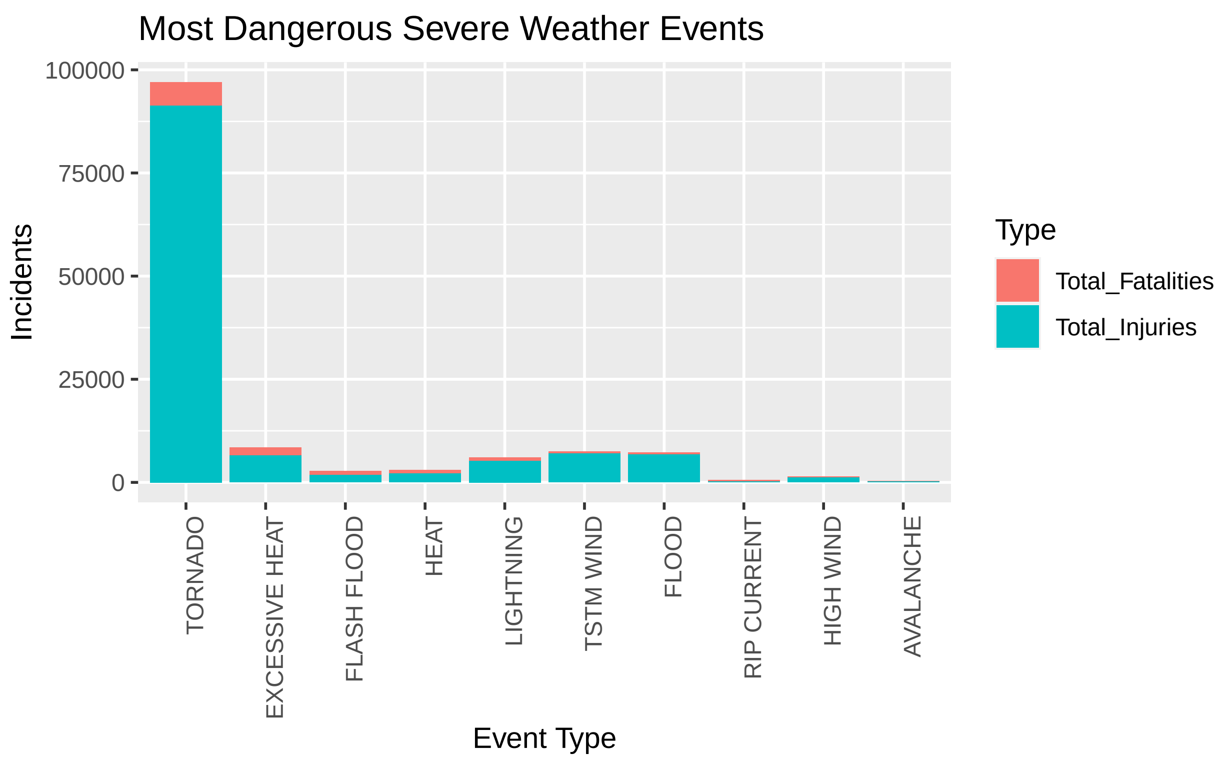
```
#For later use
IncidentRank <- econ_top25 %>% pull(EVTYPE)

econ_top25p <- econ_top25 %>% pivot_longer(Total_Prop_Damage:Total_Crop_Damage,names_to = "Type",values_

econ_top25p <- econ_top25p %>% mutate(EVTYPE = factor(EVTYPE,levels=IncidentRank))

#Convert to Millions
econ_top25p <- econ_top25p %>% mutate(Value = Value/100000)
```

Results

This section briefly presents the most dangerous and costly severe weather events based on the data set.

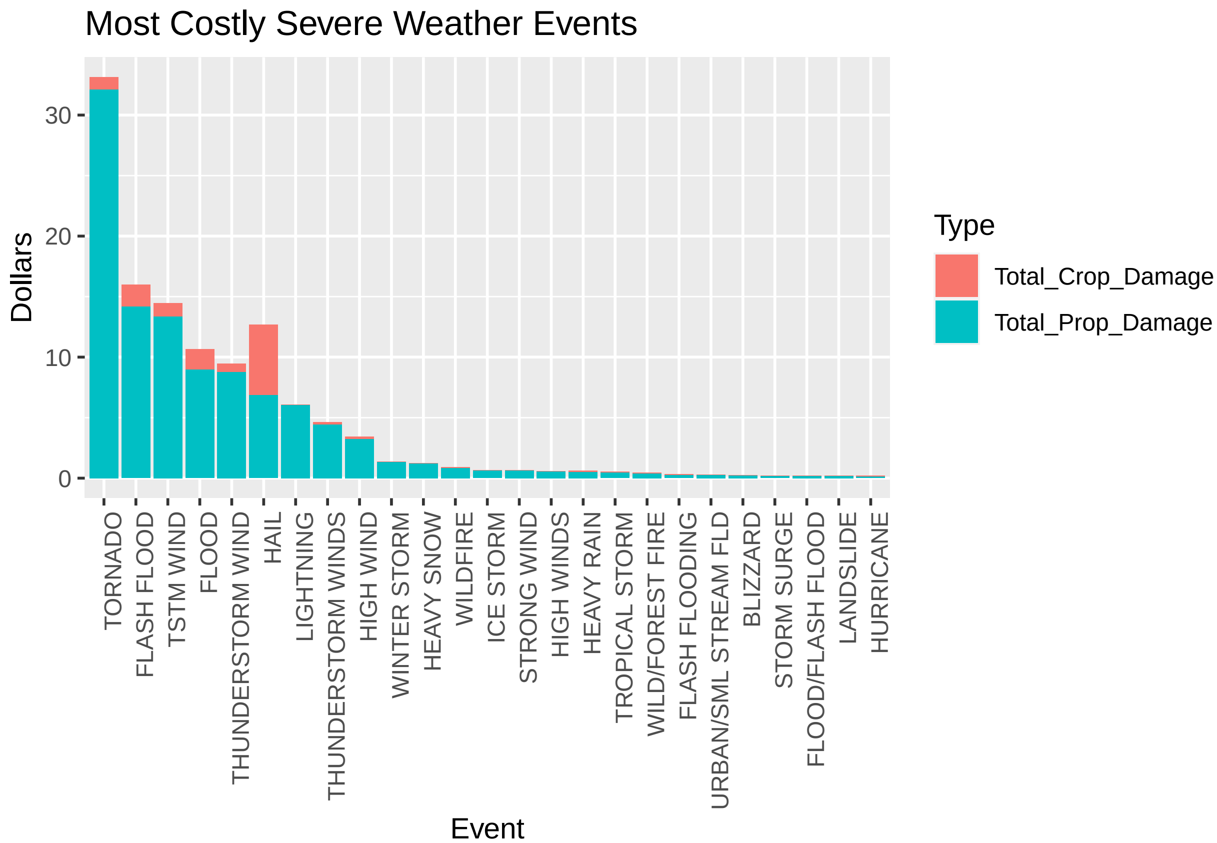
Most Dangerous Disasters

The following graph shows the 10 most dangerous disasters in terms of injuries and fatalities. The graph is ordered from left to right starting with the highest fatality rate.



Most Costly Disasters

The following graph shows the 25 most costly disasters in terms of property and crop damage. The graph is ordered from left to rate starting with the most property damage.



Conclusion

This analysis addresses the question of which types of events are most harmful to population health, and which types of events have the greatest economic consequences?

Based on these results we can conclusively state that **tornadoes** are the most serious cause of concern in both counts, dwarving many other incident types in terms of severity.

In terms of economic damage, floods (including Flash floods) and wind damage are also of serious concern.