

Health and Economic Storm Damage

Bryan Willauer

4/18/2020

Synopsis

Extreme weather events can have considerable effects on public health and economic well being. In this analysis I am going to use data from the U.S. National Oceanic and Atmospheric Administration's (NOAA) storm database to determine which weather types are the top contributors to fatalities, injuries, and economic cost.

Data Analysis

The data for this analysis was obtained from the file `repdata_data_StormData.csv.bz2`. This file was provided and contains weather events starting in the year 1950 and ending in November 2011. The data from the file was loaded into R using the following code.

```
library(tidyverse)
library(lubridate)
library(scales)
```

```
df <- read.csv("repdata_data_StormData.csv.bz2")
```

Upon inspection of the data, it was determined to limit the data used for this analysis to weather events after 1995. In the earlier years of the database there are generally fewer events recorded, most likely due to a lack of good records. More recent years should be considered more complete.

The data was separated using the following code.

```
df$year <- year(as.Date(df$BGN_DATE, format = "%m/%d/%Y"))
df1 <- df %>% filter(year >= 1995)
```

In the data, property and crop damage were listed in multiple columns with the first being a number and the second being a character that denoted a multiplier for the number given. The following code was used to calculate the actual damage using the number and multiplier columns.

```
real_damage <- function(number, multiplier){
  if(multiplier == 'B'){
    return(number * 1000000000)
  } else if(multiplier == 'M'){
    return(number * 1000000)
  } else if(multiplier == 'K'){
    return(number * 1000)
  }
}
```

```

    } else{
      return(number)
    }
  }

df1$prop_damage <- mapply(real_damage, df1$PROPDMG, df1$PROPDMGEXP)
df1$crop_damage <- mapply(real_damage, df1$CROPDMG, df1$CROPDMGEXP)
df1$econ_damage <- df1$prop_damage + df1$crop_damage

```

Results

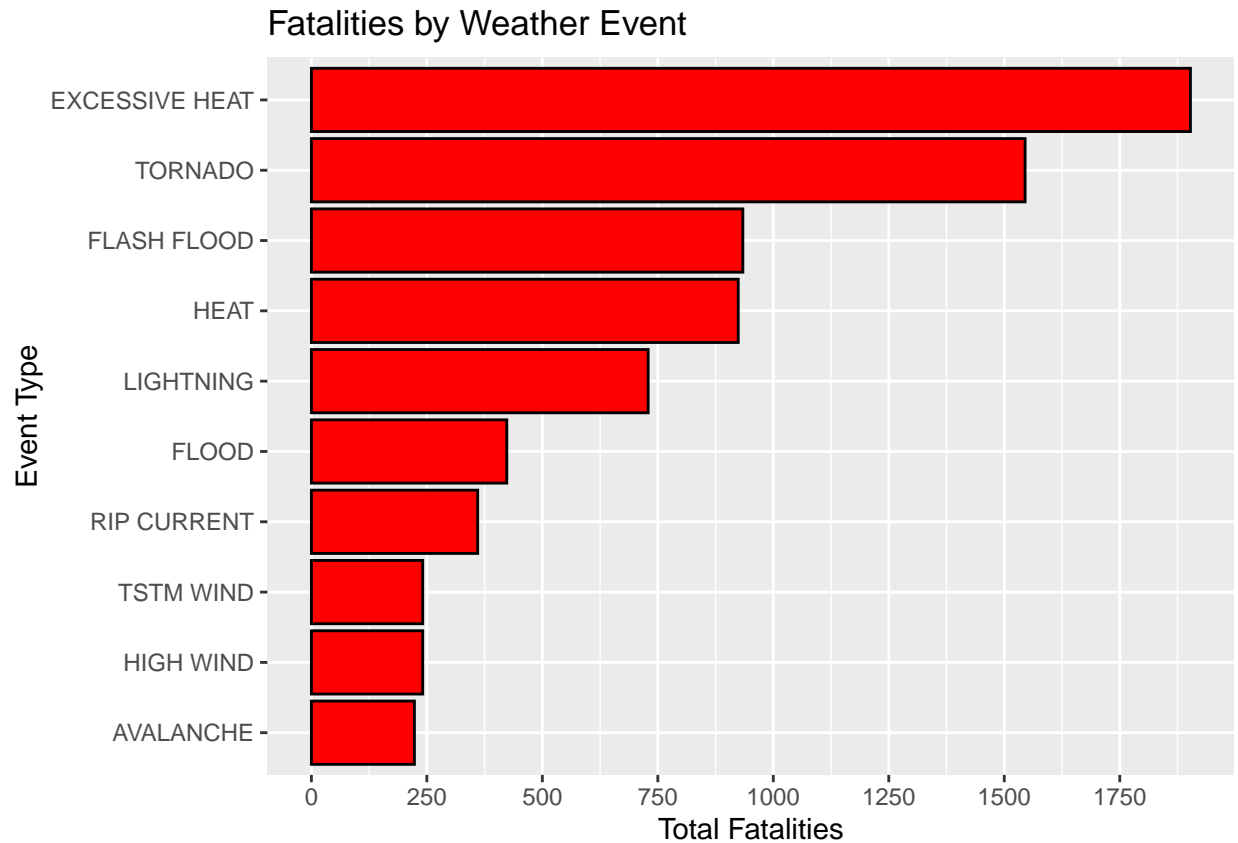
Based on the analysis, excessive heat was the top producer of fatalities between 1995 and 2011. The following graph shows the top ten fatality causing weather types during the analysis period.

```

deaths <- df1 %>% group_by(EVTYPE) %>%
  summarize(Fatalites = sum(FATALITIES)) %>%
  arrange(desc(Fatalites)) %>%
  top_n(10)

deaths %>% ggplot(aes(x = reorder(EVTYPE, Fatalites), Fatalites)) +
  geom_col(col = "Black", fill = "Red") +
  coord_flip() +
  ylab("Total Fatalities") +
  xlab("Event Type") +
  ggtitle("Fatalities by Weather Event") +
  scale_y_continuous(breaks = seq(0, 2000, len = 9))

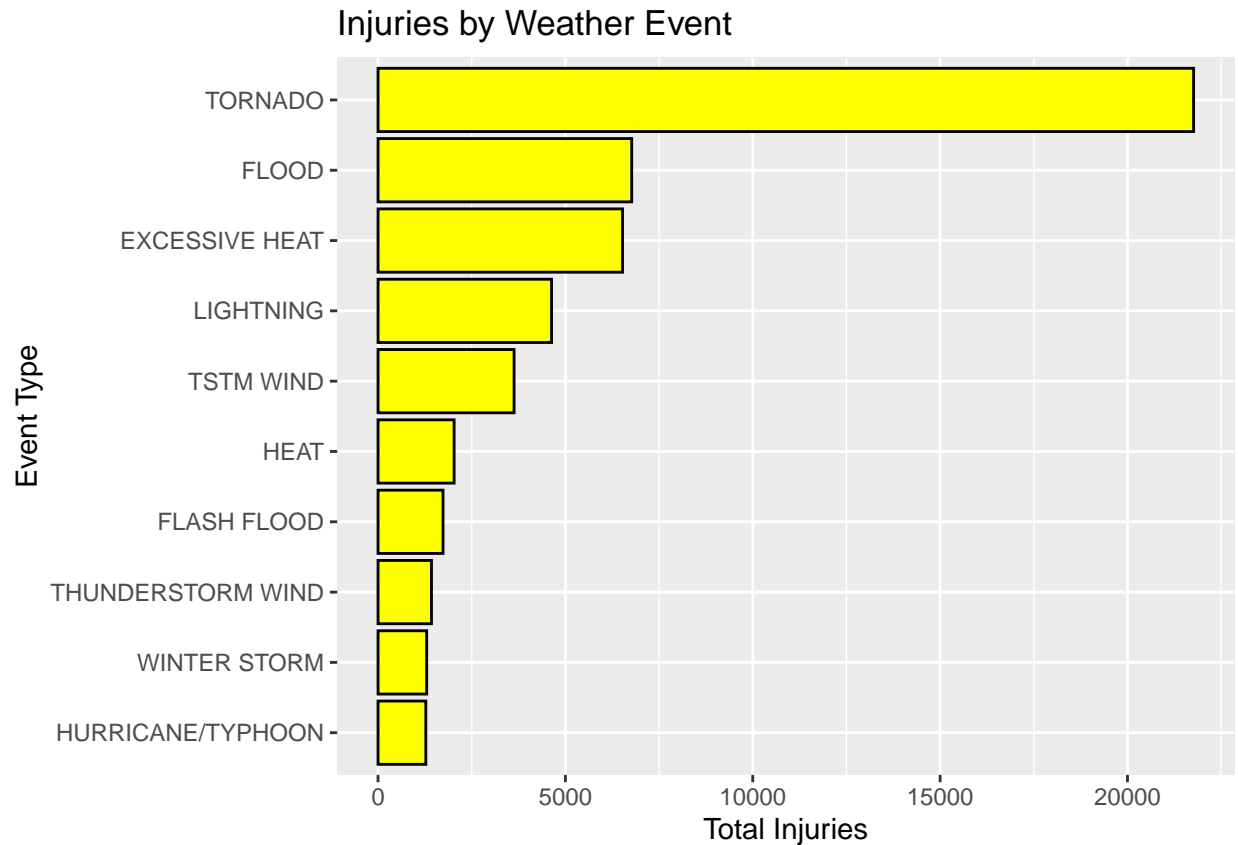
```



Tornadoes were the leading cause of injuries during the analysis period. The graph below shows the top 10 injury inducing weather types.

```
injured <- df1 %>% group_by(EVTYPE) %>%
  summarize(Injuries = sum(INJURIES)) %>%
  arrange(desc(Injuries)) %>%
  top_n(10)

injured %>% ggplot(aes(x = reorder(EVTYPE, Injuries), Injuries)) +
  geom_col(col = "Black", fill = "Yellow") +
  coord_flip() +
  ylab("Total Injuries") +
  xlab("Event Type") +
  ggtitle("Injuries by Weather Event") +
  scale_y_continuous(breaks = seq(0, 25000, len = 6))
```



Floods were the costliest weather type during the analysis period. The following graph shows the top 10 costliest weather types for the analysis period.

```
damage <- df1 %>% group_by(EVTYPE) %>%
  summarize(Total.Damage = sum(econ_damage)) %>%
  arrange(desc(Total.Damage)) %>%
  top_n(10)

damage %>% ggplot(aes(x = reorder(EVTYPE, Total.Damage), Total.Damage)) +
  geom_col(col = "Black", fill = "Orange") +
  coord_flip() +
  ylab("Total Damage (Billions of Dollars)") +
  xlab("Event Type") +
  ggtitle("Economic Damage by Weather Event") +
  scale_y_continuous(breaks = seq(0, 150000000000, len = 7), labels = unit_format(unit="", scale =
```

