Human health and economic consequences of extreme weather events

Callen Hyland

3/8/2022

Synopsis

Extreme weather events can have devastating consequences for human life and economic activity. Here we examine the economic and human cost of extreme weather events recorded by the National Weather Service between 1996 and 2011 to determine which types of events are the most damaging. Number of injuries and number of fatalities were used as measures of the human health cost, and monetary cost of property and crops destroyed were measures of economic impact. Tornadoes caused the highest number of both injuries, but excessive heat caused the greatest number of fatalities. Drought was the leading cause of crop damage, while floods were the leading cause of property damage.

Data Processing

```
library(ggplot2)
library(data.table)
library(dplyr)
library(lubridate)
library(stringdist)
library(stringr)
```

The following code chunk reads the compressed data set into a data table (902297 records on 37 variables) and reads a list of the exponential codes. The storm events data was collected by the National Weather Service and downloaded from the National Climate Data Center. The exponential codes were recorded from [this website] (https://rstudio-pubs-static.s3.amazonaws.com/58957_37b6723ee52b455990e149edde45e5b6.html).

```
# Read storm data table and exponential column codes
storm <- fread("repdata_data_StormData.csv.bz2")
exp.codes <- read.csv("storm_exp_codes.csv")
events.list <- read.csv("events_list.csv", header = FALSE) %>%
  unlist() %>%
  toupper()
```

We next remove all data before 1996 because this is the beginning of the complete records, because including data from before this time would bias the records toward the types of events that were recorded. Records with no fatalities, injuries, property damage, or crop damage are also removed.

```
# Get just the year from date
storm$BGN_DATE <- parse_date_time(storm$BGN_DATE,
                                  orders = "mdy HMS")
storm$YEAR <- year(storm$BGN_DATE)

# Select only relevant columns and remove all records before 1996
storm <- storm %>%
  select(YEAR, EVTYPE, FATALITIES, INJURIES, PROPDMG,PROPDMGEXP, CROPDMG, CROPDMGEXP) %>%
  filter(YEAR >= 1996)
```

```r
# Keep all rows that have fatalities or injuries, or have property damage or crop damage
storm <- storm %>% filter(FATALITIES != 0 |
                             INJURIES != 0 |
                             CROPDMG != 0 |
                             PROPDMG != 0)

# Make all EVTYPE descriptors uppercase and remove leading and trailing white space
storm$EVTYPE <- storm$EVTYPE %>%
  toupper() %>%
  trimws()
```

The following code matches the entered event type descriptors to the official event type names (since there are many typos). After doing these replacements, there are only 1904 records (out of an original 902297) that do not have an EVTYPE match is the official list. Most of these either have no obvious match in the official list, or account for only a small number of observations, and likely don't contribute much to the total damage and human health cost.

```r
# Replace all "TSTM" with "THUNDERSTORM"
storm$EVTYPE <- str_replace_all(storm$EVTYPE, "TSTM", "THUNDERSTORM")

# If it contains "THUNDERSTORM WIND", replace with just this string
storm$EVTYPE <- ifelse(grepl("THUNDERSTORM WIND", storm$EVTYPE), "THUNDERSTORM WIND", storm$EVTYPE)

# If it contains "WINTER", replace with WINTER WEATHER"
storm$EVTYPE <- ifelse(grepl("WINTER", storm$EVTYPE), "WINTER WEATHER", storm$EVTYPE)

# If it contains "FLD" or "FLOOD", replace with "FLOOD"
storm$EVTYPE <- ifelse(grepl("FLD", storm$EVTYPE) | grepl("FLOOD", storm$EVTYPE), "FLOOD", storm$EVTYPE)

# If it contains "HURRICANE", replace with "HURRICANE (TYPHOON)"
storm$EVTYPE <- ifelse(grepl("HURRICANE", storm$EVTYPE), "HURRICANE (TYPHOON)", storm$EVTYPE)

# If it contains "FIRE", replace with "WILDFIRE"
storm$EVTYPE <- ifelse(grepl("FIRE", storm$EVTYPE), "WILDFIRE", storm$EVTYPE)

# Match the events list with with the event list, if there is no match, remove the record
storm$match <- amatch(storm$EVTYPE, events.list)
```

The following code chunk calculates the dollar value of reported crop damage and property damage by multiplying it by an exponential factor.

```r
# Create data frames with multiplier to merge with storm
crop.codes <- data.frame(exp.codes$Exp, exp.codes$Mult)
names(crop.codes) <- c("CROPDMGEXP", "Mult.crop")
prop.codes <- data.frame(exp.codes$Exp, exp.codes$Mult)
names(prop.codes) <- c("PROPDMGEXP", "Mult.prop")

# merge multiplier with storm data set
storm <- left_join(storm, prop.codes, by = "PROPDMGEXP")
storm <- left_join(storm, crop.codes, by = "CROPDMGEXP")

# Create new columns with monetary value of crop and property damage
storm <- mutate(storm, PROPDMG.USD = Mult.prop * PROPDMG)
storm <- mutate(storm, CROPDMG.USD = Mult.crop * CROPDMG)
```

```r
# Select columns of interest
storm.sel <- storm %>% select(EVTYPE, FATALITIES, INJURIES,
                              PROPDMG.USD, CROPDMG.USD)
# remove large data sets
rm(storm, exp.codes, prop.codes, crop.codes)
```

The following code chunk creates a summary table with the total fatalities, total injuries, total crop damage, and total property damage for each type of weather event.

```r
# create a new table with sums of all columns
sum.by.event <- storm.sel %>%
  group_by(EVTYPE) %>%
  summarize(across(everything(), list(sum)))

# New names for columns
colnames(sum.by.event) <- c("Event", "Fatalities", "Injuries",
                            "Property_damage", "Crop_damage")
```

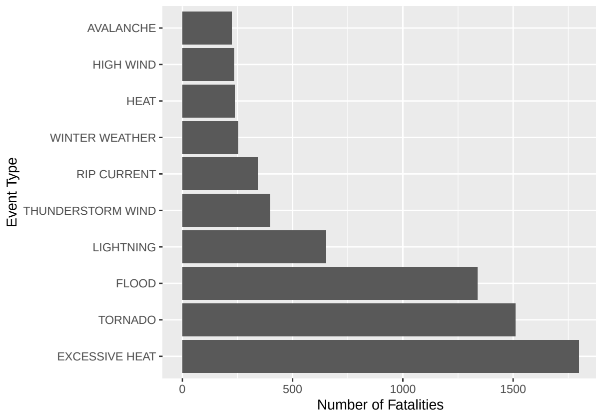
The following code chunk sorts the table to find top ten weather events associated with fatalities, injuries, property damage, and crop damage.

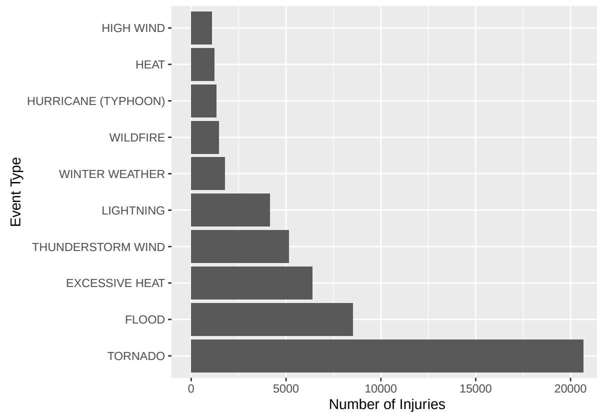
```r
fatal.top10 <- sum.by.event[order(
  sum.by.event$Fatalities, decreasing = TRUE),][1:10,c("Event", "Fatalities")]
fatal.top10$Event <- factor(fatal.top10$Event, levels = fatal.top10$Event)

injur.top10 <- sum.by.event[order(
  sum.by.event$Injuries, decreasing = TRUE),][1:10,c("Event", "Injuries")]
injur.top10$Event <- factor(injur.top10$Event, levels = injur.top10$Event)

prop.top10 <- sum.by.event[order(
  sum.by.event$Property_damage, decreasing = TRUE),][1:10,c("Event", "Property_damage")]
prop.top10$Event <- factor(prop.top10$Event, levels = prop.top10$Event)

crop.top10 <- sum.by.event[order(sum.by.event$Crop_damage,
                                 decreasing = TRUE),][1:10,c("Event", "Crop_damage")]
crop.top10$Event <- factor(crop.top10$Event, levels = crop.top10$Event)
```

Results

The figure below shows the top ten extreme weather events associated with fatalities and injuries.

```r
# subplot 1
ggplot(data = fatal.top10) +
  geom_col(aes(x = Event, y = Fatalities)) +
  coord_flip() +
  ylab("Number of Fatalities") +
  xlab("Event Type")

# subplot 2
ggplot(data = injur.top10) +
  geom_col(aes(x = Event, y = Injuries)) +
  coord_flip() +
  ylab("Number of Injuries") +
  xlab("Event Type")
```






```r
# subplot 1
ggplot(data = prop.top10) +
  geom_col(aes(x = Event, y = Property_damage)) +
  coord_flip() +
  ylab("Property Damage (USD)") +
  xlab("Event Type")

# subplot 2
ggplot(data = crop.top10) +
  geom_col(aes(x = Event, y = Crop_damage)) +
  coord_flip() +
  ylab("Crop Damage (USD)") +
  xlab("Event Type")
```