In the United States, tornadoes are the most damaging events to the health of the population and floods have the greatest economic consequences

Denize Palmito

25 de abril de 2021

1. Synopsis

This task analyzes will address climate issues that have occurred in the United States addressing issues related to public health and furthermore what implications these events have had on economic issues.

Thus, the US National Oceanic and Atmospheric Administration (NOAA) storm database will be used in this activity. This data set records the characteristics that are related to: - major storms and, - weather events, describing how and when they occur, as well as estimates of any fatalities, injuries and material damage.

Therefore, the main objective of this analysis is to avoid these negative results as far as possible.

In this analysis, two questions must be answered:

1) In the United States, what types of events (as indicated in the EVTYPE variable) are most harmful to the health of the population?

2) In the United States, what types of events have the greatest economic consequences?

The steps developed in this activity are presented below.

2. Data Processing

2.1 Data Loading

```
input <- read.csv(file = "repdata_data_StormData.csv", sep = ",",
                  stringsAsFactors = FALSE)
input$BGN_DATE <- as.Date(input$BGN_DATE, "%m/%d/%Y %H:%M:%S")
```

2.2 Data Selection

The events in the bank started in the year 1950 and end in November 2011.

```
summary(input$BGN_DATE)
```

```
##         Min.      1st Qu.       Median         Mean      3rd Qu.         Max.
## "1950-01-03" "1995-04-20" "2002-03-18" "1998-12-27" "2007-07-28" "2011-11-30"
```

```
recentdata<-input[input$BGN_DATE>as.Date("2002-01-01", "%Y-%m-%d"),]
selectNames <- c("EVTYPE","FATALITIES","INJURIES","PROPDMG","PROPDMGEXP" ,"CROPDMG","CROPDMGEXP")
data<-select(recentdata, selectNames)
```

```
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(selectNames)` instead of `selectNames` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.
```

2.3 Regrouping EVTYPE

```
data$EVTYPE[grepl("FLOOD", data$EVTYPE, ignore.case = TRUE)] <- "FLOOD"
data$EVTYPE[grepl("TORNADO", data$EVTYPE, ignore.case = TRUE)] <- "TORNADO"
data$EVTYPE[grepl("TSTM|THUNDERSTORM", data$EVTYPE, ignore.case = TRUE)] <- "TSTM"
data$EVTYPE[grepl("TROPICAL|STORM", data$EVTYPE, ignore.case = TRUE)] <- "STORM"
data$EVTYPE[grepl("HURRICANE", data$EVTYPE, ignore.case = TRUE)] <- "HURRICANE"
data$EVTYPE[grepl("ICE|SNOW|FROST|SLEET", data$EVTYPE, ignore.case = TRUE)] <- "SNOW"
data$EVTYPE[grepl("FOG", data$EVTYPE, ignore.case = TRUE)] <- "FOG"
data$EVTYPE[grepl("COLD|WINDCHILL|FREEZE|WINTER", data$EVTYPE, ignore.case = TRUE)] <- "COLD"
data$EVTYPE[grepl("HEAT|WARM|HOT", data$EVTYPE, ignore.case = TRUE)] <- "HEAT"
data$EVTYPE[grepl("CLOUD|FUNNEL", data$EVTYPE, ignore.case = TRUE)] <- "CLOUD"
data$EVTYPE[grepl("HAIL", data$EVTYPE, ignore.case = TRUE)] <- "HAIL"
data$EVTYPE[grepl("DROUGHT|DRY", data$EVTYPE, ignore.case = TRUE)] <- "DROUGHT"
data$EVTYPE[grepl("LIGHTNING", data$EVTYPE, ignore.case = TRUE)] <- "LIGHTNING"
data$EVTYPE[grepl("FIRE", data$EVTYPE, ignore.case = TRUE)] <- "FIRE"
data$EVTYPE[grepl("RAIN|SHOWER", data$EVTYPE, ignore.case = TRUE)] <- "RAIN"
data$EVTYPE[grepl("WATERSPOUT", data$EVTYPE, ignore.case = TRUE)] <- "WATERSPOUT"
data$EVTYPE[grepl("SURF", data$EVTYPE, ignore.case = TRUE)] <- "SURF"
data$EVTYPE[grepl("CURRENT", data$EVTYPE, ignore.case = TRUE)] <- "CURRENT"
data$EVTYPE[grepl("WIND|MICROBURST", data$EVTYPE, ignore.case = TRUE)] <- "WIND"
data$EVTYPE[grepl("BLIZZARD", data$EVTYPE, ignore.case = TRUE)] <- "BLIZZARD"
data$EVTYPE[grepl("SLIDE", data$EVTYPE, ignore.case = TRUE)] <- "LANDSLIDE"
data$EVTYPE[grepl("DUST", data$EVTYPE, ignore.case = TRUE)] <- "DUST"
data$EVTYPE<-factor(data$EVTYPE)
```

2.4 Calculation of Property and Corp Damage

Property damage is indicated by two variables PROPDMG and PROPDMGEXP, same for crop damage. This step firstly converts magnitude characters in PROPDMGEXP and CROPDMGEXP to numeric values and multiplies the values with PROPDMG and CROPDMG respectively.

```
data$PROPDMGEXP<-recode(data$PROPDMGEXP,'K'=1000,'M'=1000000,'B'=1000000000,.default=1)
data$CROPDMGEXP<-recode(data$CROPDMGEXP,'K'=1000,'M'=1000000,'B'=1000000000,.default=1)
data$PROPDMGVALUE <- data$PROPDMG*data$PROPDMGEXP
data$CROPDMGVALUE <- data$CROPDMG*data$CROPDMGEXP
```

3. Result

3.1 Most harmful event types to population health

The harm of events to population health is evaluated by -fatalities and - injuries caused collectively.

The top 10 most harmful event types are illustrated below.

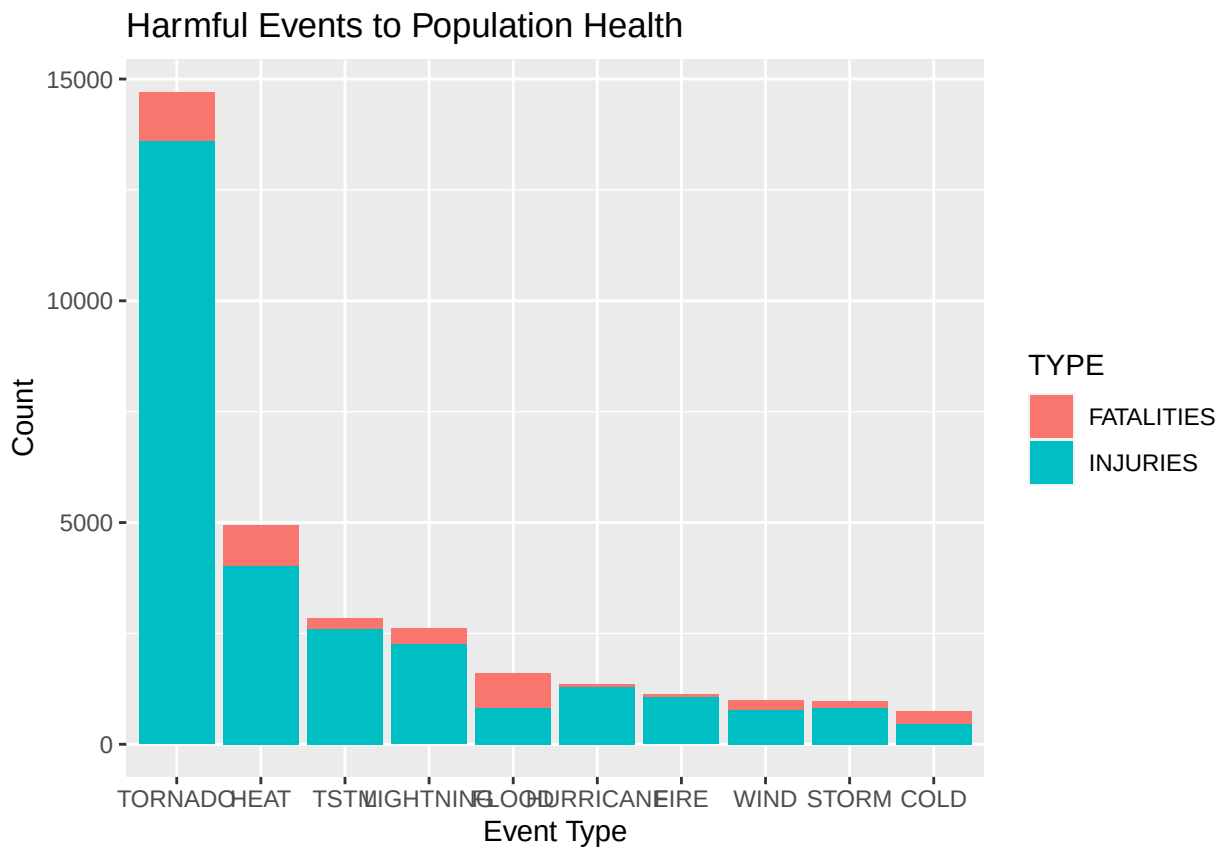
```
healthdata<-(data %>% group_by(EVTYPE) %>%
               summarise(FATALITIES = sum(FATALITIES),INJURIES = sum(INJURIES) )
               %>% arrange(desc(FATALITIES+INJURIES)))
```

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

```
mostHarm<-healthdata[1:10,]
print(mostHarm)
```

```
## # A tibble: 10 x 3
##    EVTYPE    FATALITIES INJURIES
##    <fct>          <dbl>    <dbl>
##  1 TORNADO         1112    13588
##  2 HEAT             920     4019
##  3 TSTM             227     2599
##  4 LIGHTNING        370     2250
##  5 FLOOD            789      820
##  6 HURRICANE         67     1291
##  7 FIRE              76     1051
##  8 WIND             217      767
##  9 STORM            139      823
## 10 COLD             297      454
```

```
plotdata<-gather(mostHarm, TYPE, VALUE, FATALITIES:INJURIES)
ggplot(plotdata, aes(x=reorder(EVTYPE,-VALUE), y=VALUE, fill=TYPE))+
geom_bar(stat="identity")+labs(title="Harmful Events to Population Health",
                               x="Event Type", y="Count")
```



3.2 Event types with the greatest economic consequences

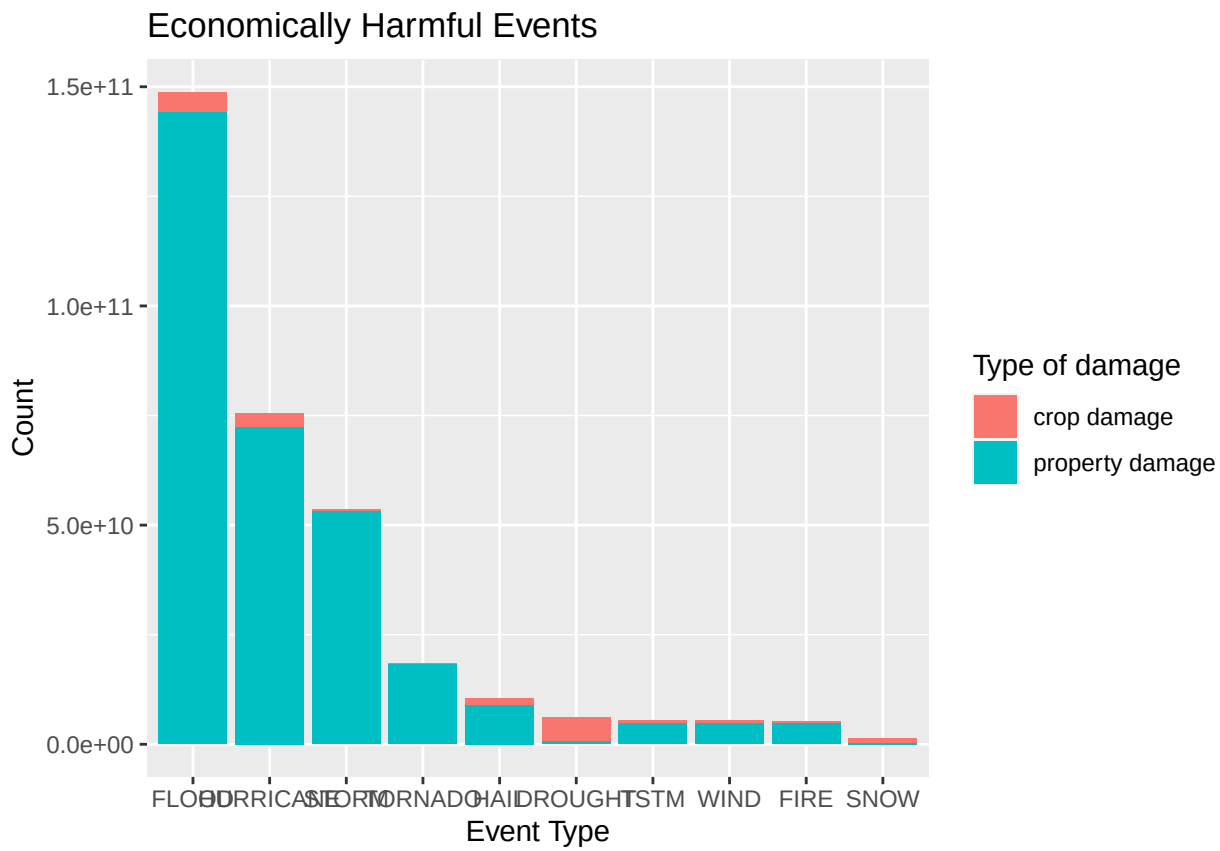
The greatest economic consequences are evaluated by -property and -crop damage collectively.

The top 10 event types with greatest economic consequences are shown below.

```
econdata<-(data %>% group_by(EVTYPE) %>% summarise(PROPDMGVALUE =
sum(PROPDMGVALUE),CROPDMGVALUE = sum(CROPDMGVALUE) ) %>%
arrange(desc(PROPDMGVALUE+CROPDMGVALUE)))
```

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

```
mostEcon<-econdata[1:10,]
plotdata2<-gather(mostEcon, TYPE, VALUE, PROPDMGVALUE:CROPDMGVALUE)
ggplot(plotdata2, aes(x=reorder(EVTYPE,-VALUE),
                      y=VALUE, fill=factor(TYPE,
                      labels=c("crop damage", "property damage"))))+
  geom_bar(stat="identity")+labs(title="Economically Harmful Events",
x="Event Type", y="Count")+guides(fill = guide_legend(title = "Type of damage"))
```



4. General Conclusion

By analyzing the data, we can conclude that in the United States, tornadoes are the most damaging event to the health of the population; Across the United States, floods have the greatest economic consequences.