Summary of Weather Event Impact on Economics and Heath

by willeslewis

Synopsis

In this report I conduct a brief exploratory analysis to indicate which weather events are probably the most damaging in terms of population health and economic value. First in the data processing section, I load the relevant storm data, and then do some simple arithmetic operations to get the damage monetary values into appropraite units. Finally I sort the data by most expensive or detrimental to health in order to answer the posed questions.

The findings in the results section, indicate that in terms of total health (including fatalities and injury), tornados are the most damaging. In terms of economic damage (including property and crop damage) floods again are the most damaging. I discuss this in slightly more detail in the results section.

Data Processing

In this section I import the relevant data using the function fread from the data.table library, reading in only the event type, number of fatalities, number of injuries, property damage, and crop damage. I then split the data into two data tables one with health and the other with economic data. I then create a multiplier for the PROPEXP/CROPEXP data values according to the mapping discussed in https://rstudio-pubs-static.s3.amazonaws.com/58957_37b6723ee52b455990e149edde45e5b6.html and multiply the PROPDMG and CROPDMG values by their appropriate multiplier. Finally add a column to each table so that the total number of fatalities and injuries is present in one and the total value of damages in the other.Below is the code for loading and subsetting the appropriate data.

```
##load plyr
library(plyr)
library(data.table)

##Set current directory
dir<-"/Users/William/Desktop/Desktop/Projects/coursera-dsc/course-5/FinalProject"
workdir<-setwd(dir)

mydat<-fread("data/StormData.csv",
             select =
c("EVTYPE","FATALITIES","INJURIES","PROPDMG","PROPDMGEXP","CROPDMG","CROPDMGEXP"))
```

Read 88.9% of 967216 rows Read 902297 rows and 7 (of 37) columns from 0.523 GB file in 00:00:03

```
##mydat must be transformed according to the value of EXP as described above.
##First I will make all exponent values lower case to make the
##mulitiplier evaluation easier.
mydat$PROPDMGEXP<-tolower(mydat$PROPDMGEXP)

mydat$CROPDMGEXP<-tolower(mydat$CROPDMGEXP)

##Multiplier dictionary and values
dic<-c("","-","?","+","0","1","2","3","4","5","6","7","8","h","k","m","b")
```

```
dicval<-c(0,0,0,1,10,10,10,10,10,10,10,10,10,100,1000,1000000,1000000000)



##Multiply DMG values by appropriate number
mydat$PROPDMG<-mydat$PROPDMG*dicval[match(mydat$PROPDMGEXP,dic)]
mydat$CROPDMG<-mydat$CROPDMG*dicval[match(mydat$CROPDMGEXP,dic)]

##Get sum of fatalaties injuries, property and crop damage for each event.
EventHealth<-ddply(mydat,.(EVTYPE),summarize,
                   TotalFat=sum(FATALITIES),TotalInj=sum(INJURIES))

EventDMG<-ddply(mydat,.(EVTYPE),summarize,
                TotProp=sum(PROPDMG),TotCrop=sum(CROPDMG))

EventHealth$TotHealth<-EventHealth$TotalFat+EventHealth$TotalInj

EventDMG$TotDMG<-EventDMG$TotProp+EventDMG$TotCrop

OrderedHealth<-setorder(EventHealth,-TotHealth)

OrderedDMG<-setorder(EventDMG,-TotDMG)
```

Results

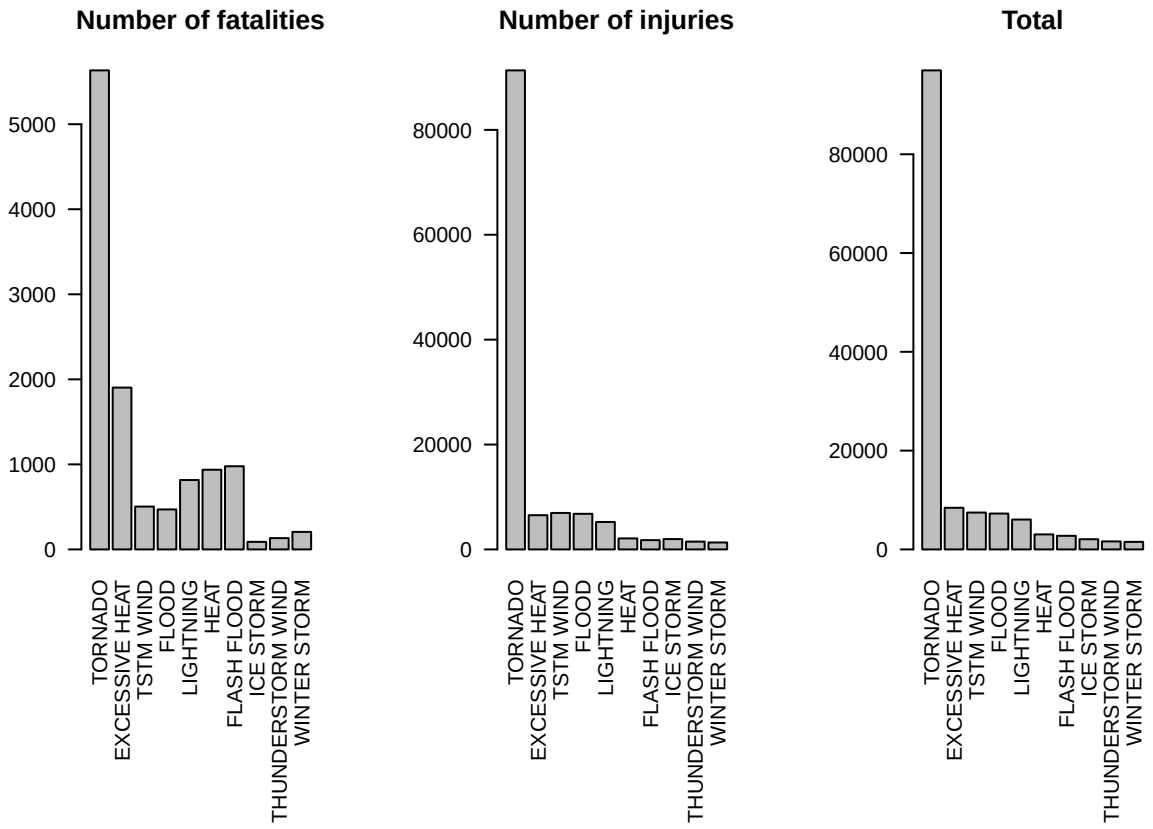
Here I plot the results of the above data processing section and highlight the main features. ### Health Effects To understand health effects, the figure below plots number of fatalities, injuries, and the sum of the two for the ten events with the highest total number of fatalities and injuries. This shows that for all three cases, tornados are the most detremental event, although the exact ordering of subsequent events is not necessarily preserved. Injuries are far more common that fatalities in weather related health effects and hence are the primary determing factor for the total number of health related effects for each event.

```
        par(mfrow=c(1,3),mar=c(11,5,4,2))

        barplot(OrderedHealth$TotalFat[1:10],names.arg=OrderedHealth$EVTYPE[1:10],
                las=2,main="Number of fatalities")

        barplot(OrderedHealth$TotalInj[1:10],names.arg=OrderedHealth$EVTYPE[1:10],
                las=2,main="Number of injuries")

        barplot(OrderedHealth$TotHealth[1:10],names.arg=OrderedHealth$EVTYPE[1:10],
                las=2,main="Total")
```



Economic Impact

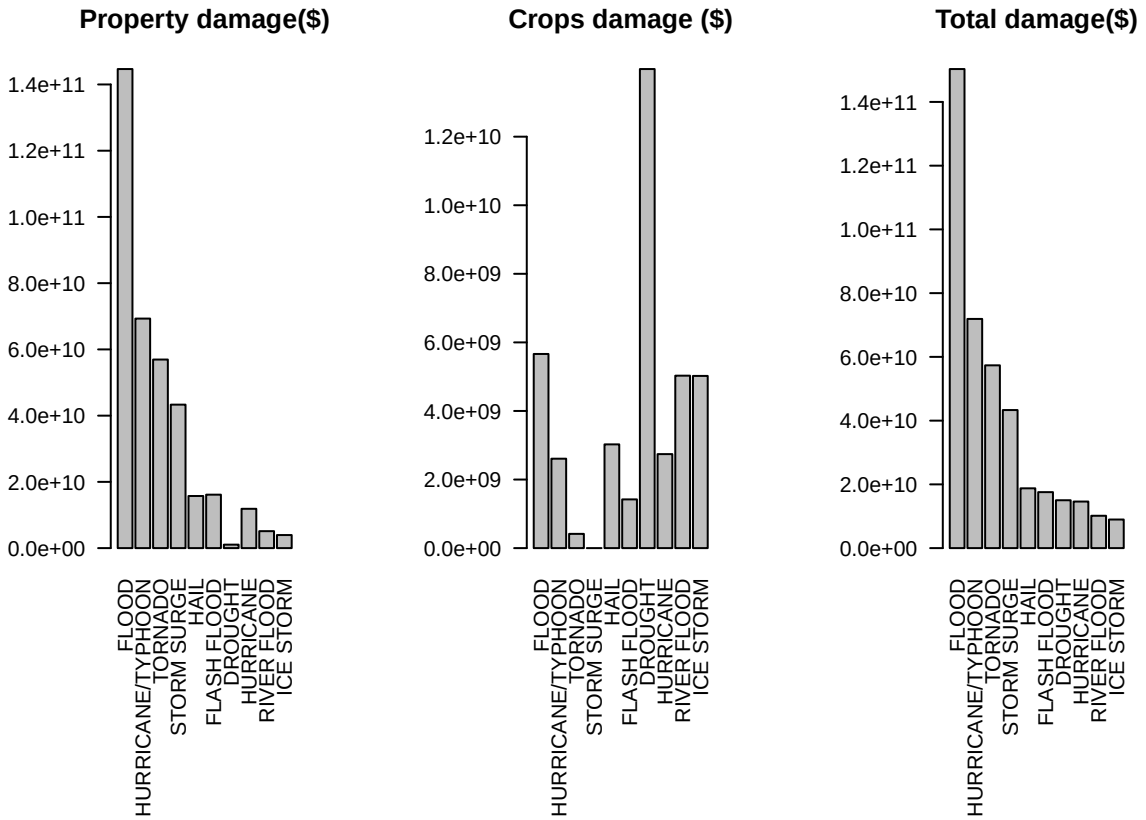
To understand economic effects, the figure below plots estimated damage (in dollars) caused by each event type to property, crops, and the sum of the two for the ten events with the highest total damage as the sum of property and crops. This shows that floods are the most detremental event, although the exact ordering is not preserved and for crops, drought is much more damaging. Property damages are typically about an order of magnitude or two more expensive and hence are the primary determing factor for the total economic damage event.

```
        par(mfrow=c(1,3),mar=c(11,7,4,2))

        barplot(OrderedDMG$TotProp[1:10],names.arg=OrderedDMG$EVTYPE[1:10],
                las=2,main="Property damage($)")

        barplot(OrderedDMG$TotCrop[1:10],names.arg=OrderedDMG$EVTYPE[1:10],
                las=2,main="Crops damage ($)")

        barplot(OrderedDMG$TotDMG[1:10],names.arg=OrderedDMG$EVTYPE[1:10],
                las=2,main="Total damage($)")
```



Recommendations for Further Study

The above analysis is quite simple and only studies the total cost across all of the USA for the events. To make actionable plans, one would want at the very least to study these effects by region as well.