The influence of weather on population health and economy

ChengWen Luo

23/03/2018

Synopsis

Storms and other severe weather events can cause both public health and economic problems for communities and municipalities. Many severe events can result in fatalities, injuries, and property damage, and preventing such outcomes to the extent possible is a key concern. This project involves exploring the U.S. National Oceanic and Atmospheric Administration's (NOAA) storm database. We asked two questions in the current assignment: (1) Across the United States, which types of events (as indicated in the ??????? variable) are most harmful with respect to population health? (2) Across the United States, which types of events have the greatest economic consequences?. The assignment has two parts: section 1 describes the data processing, and section 2 presents the main findings. The key variables used are: "FATALITIES", "INJURIES","PROPDMG","CROPDMG" and "EVTYPE".

Data processing

```
setwd("/Volumes/Daisy/R/R_assignmant/R_markdown/final/")
dat<-read.csv("repdata-data-StormData.csv.bz2",head=T)
```

sum of death caused by different events

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
frequency.event<-data.frame(tapply(dat$EVTYPE,dat$EVTYPE,length))
names(frequency.event) <- "event.frequency"
head(frequency.event,10)
```

```
##                       event.frequency
##    HIGH SURF ADVISORY                1
##  COASTAL FLOOD                       1
##  FLASH FLOOD                         1
##  LIGHTNING                           1
```

```
##  TSTM WIND                         4
##  TSTM WIND (G45)                   1
##  WATERSPOUT                        1
##  WIND                              1
## ?                                  1
## ABNORMAL WARMTH                    4
```

Many events are essentially the same type, but are considered as different categories, here I group them together.

1. check how many unique event types

```
unique.evt<-unique(dat$EVTYPE)
length(unique.evt)

## [1] 985
```

2. changing events by lowering case and excluding punctuations/blanks

```
dat$EVTYPE <- tolower (dat$EVTYPE)
dat$EVTYPE <- gsub("[[:blank:][:punct:]+]", " ", dat$EVTYPE)
dat[grepl("hurricane",dat$EVTYPE),]$EVTYPE = "huricane"
dat[grepl("severe thunderstorm",dat$EVTYPE),]$EVTYPE = "severe thunderstorm"
dat[grepl("tstm wind",dat$EVTYPE),]$EVTYPE = "tstm wind"
dat[grepl("flood",dat$EVTYPE),]$EVTYPE = "flood"
dat[grepl("heat",dat$EVTYPE),]$EVTYPE = "heat"
dat[grepl("thunderstorm",dat$EVTYPE),]$EVTYPE = "thunderstorm"
unique.evt<-unique(dat$EVTYPE)
length(unique.evt)

## [1] 662

head(unique.evt)

## [1] "tornado"       "tstm wind"     "hail"          "freezing rain"
## [5] "snow"          "flood"
```

Results

```
sum_death<-data.frame(tapply(dat$FATALITIES,dat$EVTYPE,sum,na.rm=T))
sum_death$event.type<-rownames(sum_death)
rownames(sum_death)<-1:dim(sum_death)[1]
names(sum_death)<-c("sum.death","event.type")
death.order<-sum_death[order(sum_death$sum.death ,decreasing=TRUE),]
death.order$rank.fatalities <-1:dim(sum_death)[1]
head.de<-head(death.order,6)
head.de

##     sum.death  event.type rank.fatalities
## 559      5633     tornado               1
## 189      3138        heat               2
## 120      1525       flood               3
## 328       816   lightning               4
```

```
## 579       544   tstm wind               5
## 418       368 rip current               6
```

Injury caused by different types of event

```
sum_injury<-data.frame(tapply(dat$INJURIES,dat$EVTYPE,sum,na.rm=T))
sum_injury$event.type<-rownames(sum_injury)
rownames(sum_injury)<-1:dim(sum_injury)[1]
names(sum_injury)<-c("sum.injury","event.type")
injury.order<-sum_injury[order(sum_injury$sum.injury ,decreasing=TRUE),]
injury.order$rank.INJURIES <-1:dim(sum_injury)[1]
head.in<-head(injury.order,6)
head.in
```

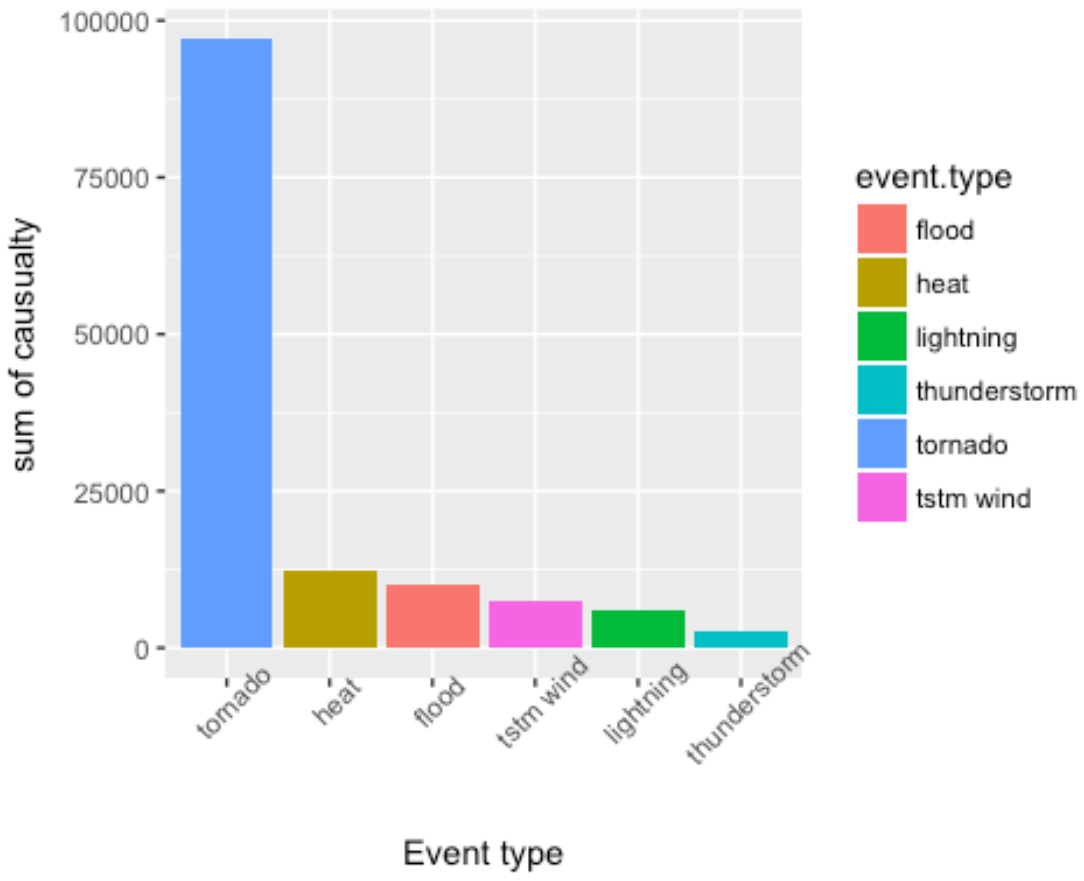
```
##     sum.injury   event.type rank.INJURIES
## 559      91346      tornado             1
## 189       9224         heat             2
## 120       8604        flood             3
## 579       7065    tstm wind             4
## 328       5230    lightning             5
## 552       2479 thunderstorm             6
```

Casualty rate caused by different types of event

```
dat$CASUALTY <- dat$INJURIES + dat$FATALITIES
sum_casualty<-data.frame(tapply(dat$CASUALTY,dat$EVTYPE,sum,na.rm=T))
sum_casualty$event.type<-rownames(sum_casualty)
rownames(sum_casualty)<-1:dim(sum_casualty)[1]
names(sum_casualty)<-c("sum.casualty","event.type")
casualty.order<-sum_casualty[order(sum_casualty$sum.casualty
,decreasing=TRUE),]
casualty.order$rank.CASUALTY <-1:dim(sum_casualty)[1]
top.casualty<-head(casualty.order,6)
top.event<-top.casualty$event.type
top.casualty
```

```
##     sum.casualty   event.type rank.CASUALTY
## 559        96979      tornado             1
## 189        12362         heat             2
## 120        10129        flood             3
## 579         7609    tstm wind             4
## 328         6046    lightning             5
## 552         2689 thunderstorm             6
```

```
library(ggplot2)
m<-ggplot(top.casualty,aes(x=reorder(event.type,
rank.CASUALTY),y=sum.casualty,fill=event.type))
m + geom_col() + theme(axis.text.x =  (element_text(angle=45))) + xlab('Event
type') + ylab ("sum of causualty")
```



Sum of property damage due to various events

```
sum.property.damage<-data.frame(tapply(dat$PROPDMG,dat$EVTYPE,sum,na.rm=T))
sum.property.damage$event.type<-rownames(sum.property.damage)
rownames(sum.property.damage)<-1:dim(sum.property.damage)[1]
names(sum.property.damage)<-c("sum.property","event.type")
property.order<-sum.property.damage[order(sum.property.damage$sum.property
,decreasing=TRUE),]
property.order$rank.property <-1:dim(sum.property.damage)[1]
top.property<-head(property.order,6)
top.event<-top.property$event.type
top.property
```

```
##     sum.property   event.type rank.property
## 559    3212258.2      tornado             1
## 120    2436131.5        flood             2
## 579    1347464.3    tstm wind             3
## 552    1329707.6 thunderstorm             4
## 159     688693.4         hail             5
## 328     603351.8    lightning             6
```

Sum of crop damage due to various events

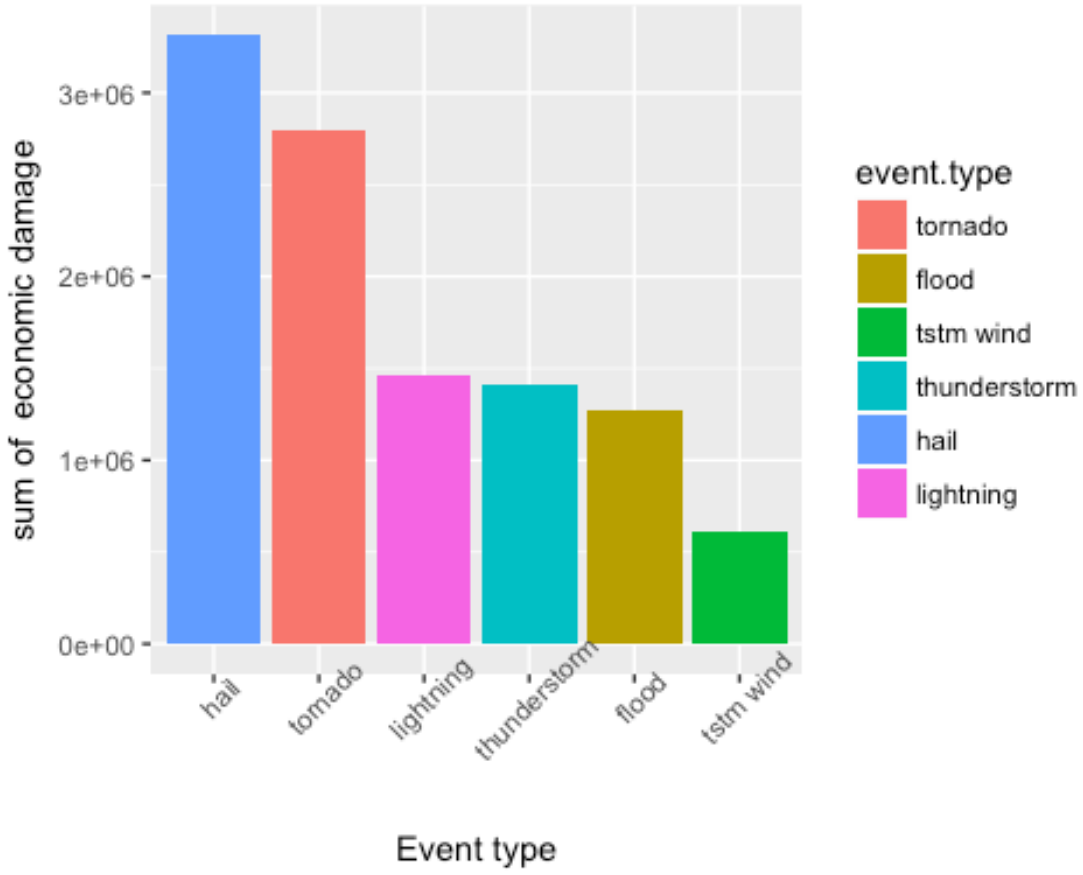
```
sum.crop.damage<-data.frame(tapply(dat$CROPDMG,dat$EVTYPE,sum,na.rm=T))
sum.crop.damage$event.type<-rownames(sum.crop.damage)
rownames(sum.crop.damage)<-1:dim(sum.crop.damage)[1]
names(sum.crop.damage)<-c("sum.crop","event.type")
crop.order<-sum.crop.damage[order(sum.crop.damage$sum.crop
,decreasing=TRUE),]
crop.order$rank.crop <-1:dim(sum.crop.damage)[1]
top.crop<-head(crop.order,6)
top.event<-top.crop$event.type
top.crop
```

```
##      sum.crop   event.type rank.crop
## 159 579596.28         hail         1
## 120 364506.73        flood         2
## 579 113561.75    tstm wind         3
## 559 100018.52      tornado         4
## 552  85801.13 thunderstorm         5
## 70   33898.62      drought         6
```

Sum of total damage due to various events

```
dat$ECO <- dat$CROPDMG + dat$PROPDMG
sum.economic.damage<-data.frame(tapply(dat$ECO,dat$EVTYPE,sum,na.rm=T))
sum.economic.damage$event.type<-rownames(sum.economic.damage)
rownames(sum.economic.damage)<-1:dim(sum.economic.damage)[1]
names(sum.economic.damage)<-c("sum.economic","event.type")
economic.order<-sum.economic.damage[order(sum.economic.damage$sum.economic
,decreasing=TRUE),]
economic.order$rank.economic <-1:dim(sum.economic.damage)[1]
top.economic<-head(economic.order,6)
top.economic$event.type <- as.factor(top.economic$event.type)
levels(top.economic$event.type) <-
top.economic[top.economic$rank.economic[1:6],]$event.type

g<-
ggplot(top.economic,aes(x=reorder(event.type,rank.economic),y=sum.economic,fi
ll=event.type))+ geom_col() + theme(axis.text.x = element_text(angle=45)) +
xlab('Event type') + ylab ("sum of  economic damage")
g
```



Conclusion

The results show that tornado is the most severe cause of casualty, following by heat, flood, tstm wind, lightening and ice storm. As to which types of natural disasters lead to greater economic damages, our analyses indicate that hail is the leading cause, followed by tornado, lightening, thunderstorm, flood and tstm wind.