RR - Week 4 - Course Project 2

Hitesh Kumar

January 9, 2018

EDA on Storm Data

Synopsis

The report presented uses data from the U.S. National Oceanic and Atmospheric Administration's (NOAA) storm database. Using this data, we will analyze the impact of severe weather events on population health and economics. The goal will be to answer which type of events are most harmful to population health and also the most damaging as it relates to economic consequences.

Data Processing

Set the working directory and load support libraries we'll use.

```
 setwd("/home/woodman/")
storm <- read.csv("repdata_data_StormData.csv", header = T)
```

Filtering the releveant columns

```
names(storm)
```

```
##  [1] "STATE__"    "BGN_DATE"   "BGN_TIME"   "TIME_ZONE"  "COUNTY"
##  [6] "COUNTYNAME" "STATE"      "EVTYPE"     "BGN_RANGE"  "BGN_AZI"
## [11] "BGN_LOCATI" "END_DATE"   "END_TIME"   "COUNTY_END" "COUNTYENDN"
## [16] "END_RANGE"  "END_AZI"    "END_LOCATI" "LENGTH"     "WIDTH"
## [21] "F"          "MAG"        "FATALITIES" "INJURIES"   "PROPDMG"
## [26] "PROPDMGEXP" "CROPDMG"    "CROPDMGEXP" "WFO"        "STATEOFFIC"
## [31] "ZONENAMES"  "LATITUDE"   "LONGITUDE"  "LATITUDE_E" "LONGITUDE_"
## [36] "REMARKS"    "REFNUM"
```

```
health<-storm[,c(8,23:24)]
property<-storm[,c(8,25:28)]
```

Result

Harmful Events with respect to Population Health

With respect to the Population Health, there are two damages caused : fatalities and injuries. The top 10 severity of the harmful events are analysed and plotted below.

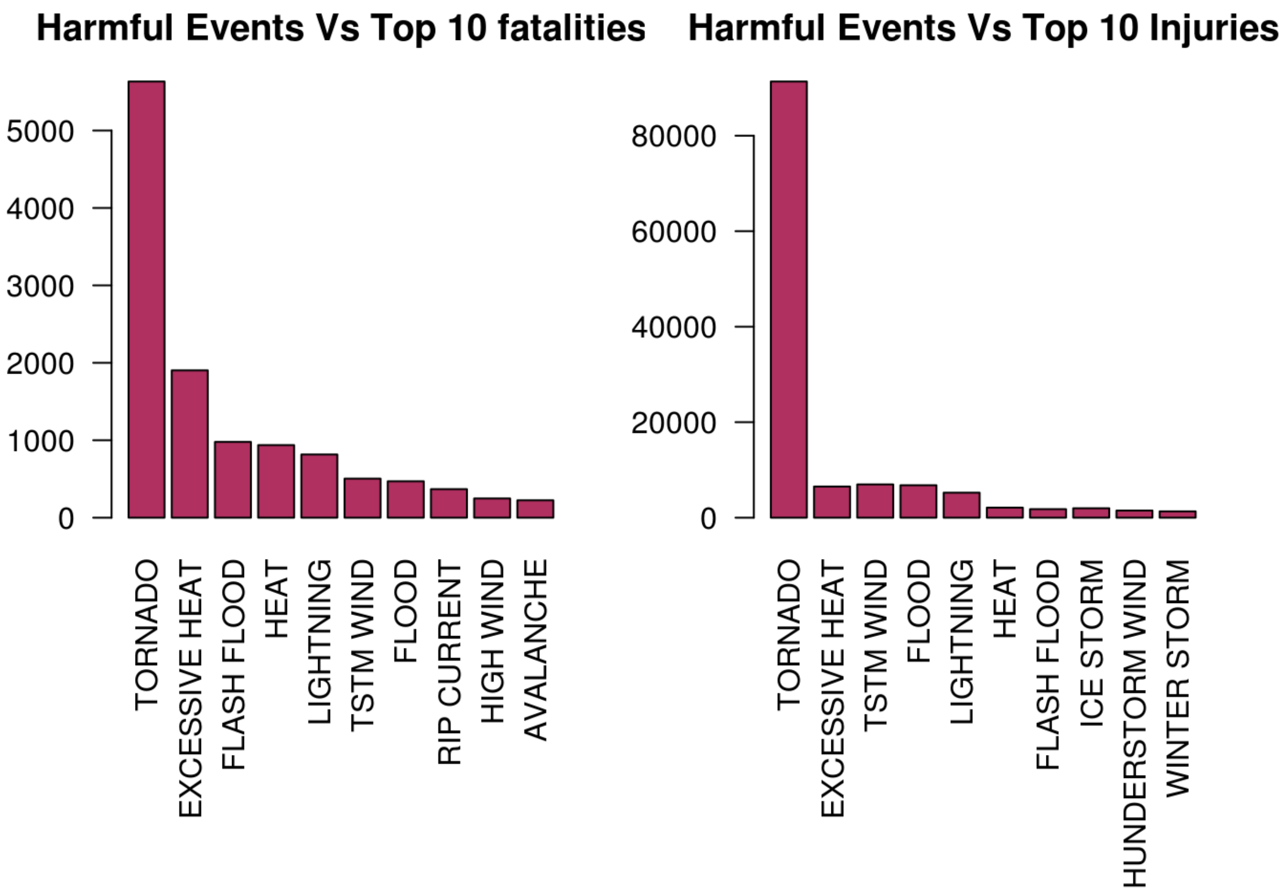
```
TOPhealth<-aggregate(cbind(FATALITIES,INJURIES) ~ EVTYPE, data = health, sum,
na.rm=TRUE)
TOPhealth<-arrange(TOPhealth, desc(FATALITIES+INJURIES))
TOPhealth<-TOPhealth[1:10,]
TOPhealth
```

```
##                EVTYPE FATALITIES INJURIES
## 1             TORNADO       5633    91346
## 2      EXCESSIVE HEAT       1903     6525
## 3           TSTM WIND        504     6957
## 4               FLOOD        470     6789
## 5           LIGHTNING        816     5230
## 6                HEAT        937     2100
## 7         FLASH FLOOD        978     1777
## 8           ICE STORM         89     1975
## 9   THUNDERSTORM WIND        133     1488
## 10       WINTER STORM        206     1321
```

```
fatal <- aggregate(FATALITIES~EVTYPE,data=storm,FUN=sum,na.rm=TRUE)
fatal <- fatal[with(fatal,order(-FATALITIES)),]
fatal <- head(fatal,10)
fatal
```

```
##             EVTYPE FATALITIES
## 830        TORNADO       5633
## 123 EXCESSIVE HEAT       1903
## 147    FLASH FLOOD        978
## 269           HEAT        937
## 452      LIGHTNING        816
## 854      TSTM WIND        504
## 164          FLOOD        470
## 581    RIP CURRENT        368
## 354      HIGH WIND        248
## 11       AVALANCHE        224
```

```
par(mfrow=c(1,2),mar=c(10,3,3,2))
barplot(fatal$FATALITIES,names.arg=fatal$EVTYPE,las=2,col="maroon",ylab="fata
lities",main="Harmful Events Vs Top 10 fatalities")
barplot(TOPhealth$INJURIES,names.arg=TOPhealth$EVTYPE,las=2,col="maroon",ylab
="injuries",main="Harmful Events Vs Top 10 Injuries")
```



###Economic Impact

Now we wil run the code for the severe weather events impact on economics. In order to do so we must first translate the values for PROPDMG and CROPDMG based on the EXP field. In looking at the field we see some anomolies in the data.

```
table(property$PROPDMGEXP)
```

```
##
##                -       ?       +       0       1       2       3       4       5
## 465934       1       8       5     216      25      13       4       4      28
##       6       7       8       B       h       H       K       m       M
##       4       5       1      40       1       6  424665       7   11330
```

```
table(property$CROPDMGEXP)
```

```
##
##             ?      0      2      B      k      K      m      M
## 618413      7     19      1      9     21 281832      1   1994
```

To simplify the data and the analysis we will focus our calculation on the alpha designators for hundreds, thousands, etc…. Note: the count for non “blanks” other than alphanumeric designators is immaterial to the population size as a whole.

Calculating the property and crop damage based on the alpha designators.

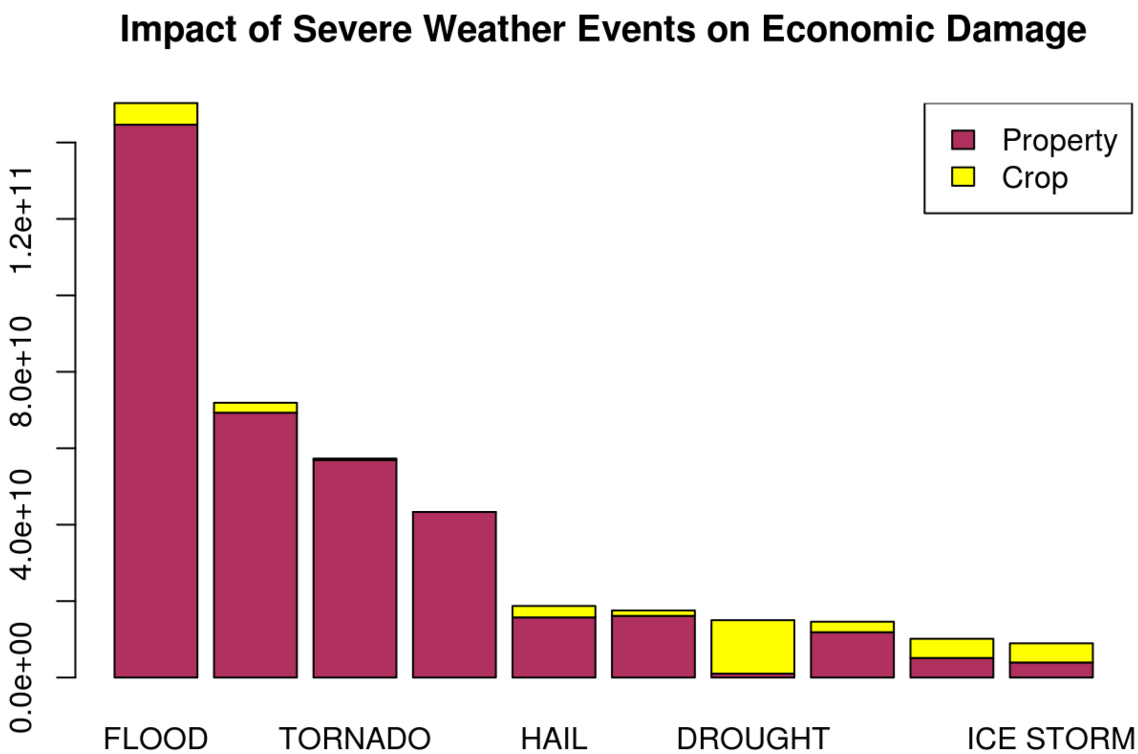
```
property$PROPDMGCALC [property$PROPDMG==0] <- 0
property$CROPDMGCALC [property$CROPDMG==0] <- 0
property$PROPDMGCALC [property$PROPDMGEXP=="H"| property$PROPDMGEXP=="h"]<- p
roperty$PROPDMG[property$PROPDMGEXP=="H"|property$PROPDMGEXP=="h"]*100
property$CROPDMGCALC [property$CROPDMGEXP=="H"| property$CROPDMGEXP=="h"]<- p
roperty$CROPDMG[property$CROPDMGEXP=="H"|property$CROPDMGEXP=="h"]*100
property$PROPDMGCALC [property$PROPDMGEXP=="K"| property$PROPDMGEXP=="k"]<- p
roperty$PROPDMG[property$PROPDMGEXP=="K"|property$PROPDMGEXP=="k"]*1000
property$CROPDMGCALC [property$CROPDMGEXP=="K"| property$CROPDMGEXP=="k"]<- p
roperty$CROPDMG[property$CROPDMGEXP=="K"|property$CROPDMGEXP=="k"]*1000
property$PROPDMGCALC [property$PROPDMGEXP=="M"| property$PROPDMGEXP=="m"]<- p
roperty$PROPDMG[property$PROPDMGEXP=="M"|property$PROPDMGEXP=="m"]*1000000
property$CROPDMGCALC [property$CROPDMGEXP=="M"| property$CROPDMGEXP=="m"]<- p
roperty$CROPDMG[property$CROPDMGEXP=="M"|property$CROPDMGEXP=="m"]*1000000
property$PROPDMGCALC [property$PROPDMGEXP=="B"| property$PROPDMGEXP=="b"]<- p
roperty$PROPDMG[property$PROPDMGEXP=="B"|property$PROPDMGEXP=="b"]*1000000000
property$CROPDMGCALC [property$CROPDMGEXP=="B"| property$CROPDMGEXP=="b"]<- p
roperty$CROPDMG[property$CROPDMGEXP=="B"|property$CROPDMGEXP=="b"]*1000000000
```

```
damage<- aggregate(cbind(PROPDMGCALC,CROPDMGCALC)~EVTYPE, data = property, su
m, na.rm=TRUE)
damage<- arrange(damage, desc(PROPDMGCALC+CROPDMGCALC))
damage<- damage[1:10,]
damage
```

```
##                EVTYPE  PROPDMGCALC CROPDMGCALC
## 1               FLOOD 144657709800  5661968450
## 2  HURRICANE/TYPHOON  69305840000  2607872800
## 3             TORNADO  56936990480   364950110
## 4         STORM SURGE  43323536000        5000
## 5                HAIL  15732262220  3000949450
## 6         FLASH FLOOD  16140811510  1420717100
## 7             DROUGHT   1046106000 13972566000
## 8           HURRICANE  11868319010  2741910000
## 9         RIVER FLOOD   5118945500  5029459000
## 10          ICE STORM   3944927810  5022110000
```

Floods top the list as the most impactful to property damage

```
x <- damage$EVTYPE
damage<- as.matrix(t(damage[,-1]))
colnames(damage)<-x
barplot(damage, col = c("Maroon", "yellow"), main = "Impact of Severe Weather
Events on Economic Damage")
legend("topright", c("Property","Crop"), fill = c("Maroon", "yellow"), bty =
"x")
```



Summary

Interesting that hurricanes dominates cost damage - and the bulk of that comes from property damage. Not surprising that droughts are mostly costly to crops!