Storm

Oye Akinnikawe

December 15, 2016

Impacts of Storms and other severe weather events in the US

1. Synopsis

This is project 2 of the Cousera Reproducible Research course. The goal of the project is to explore the NOAA storm database and analyze the impacts of events types to population health and economic consequences.

The data covers events from the year 1950 to 2011. There are fewer recorded events in earlier years compared to more recent years.

The analysis aims to investigate which different types of severe weather events are most harmful on the populations health in respect of general injuries and fatalities. Further, the economic consequences will be analyzed by exploring the financial damage done to both general property and agriculture (i.e. crops)

2. Data Processing

The Storm data is a bzip2 file that was downloaded from the cousera website. The first step is to read the data.

```
data <- read.csv("StormData.csv.bz2")
```

The required analysis deals with population health and economic analysis, therefore a subset of the required columns is created

```
storm <- data[, c("EVTYPE", "FATALITIES", "INJURIES", "PROPDMG", "PROPDMGEXP",
                  "CROPDMG", "CROPDMGEXP")]
```

2.1 Population Health

Summarizing fatalities and injuries according to the event type.

```
library(plyr)
```

```
## Warning: package 'plyr' was built under R version 3.2.5
```

```
pop_affected <- ddply(storm, .(EVTYPE), summarize, fatalities = sum(FATALITIES),
                      injuries = sum(INJURIES))
fatal <- pop_affected[order(pop_affected$fatalities, decreasing = T), ]
injury <- pop_affected[order(pop_affected$injuries, decreasing = T), ]
```

2.2 Economic Cosequences

Since the exponential values are stored in a seperate column describing their value with letters (h = hundred, k = thousand, m = million, b = billion). A function that converts the letter value of the exponent to a usable number must be implemented.

```
getExp <- function(e) {
    if (e %in% c("h", "H"))
        return(2)
    else if (e %in% c("k", "K"))
        return(3)
    else if (e %in% c("m", "M"))
        return(6)
    else if (e %in% c("b", "B"))
        return(9)
    else if (!is.na(as.numeric(e)))
        return(as.numeric(e))
    else if (e %in% c("", "-", "?", "+"))
        return(0)
    else {
        stop("Invalid value.")
    }
}
```

Then, using this function, the proper values are calculated for property damage and crop damage

```
propExp <- sapply(storm$PROPDMGEXP, FUN=getExp)
storm$propDamage <- storm$PROPDMG * (10 ** propExp)
cropExp <- sapply(storm$CROPDMGEXP, FUN=getExp)
storm$cropDamage <- storm$CROPDMG * (10 ** cropExp)
```

Summarizing finacial damage for crops and property according to event type

```
econDamage <- ddply(storm, .(EVTYPE), summarize,propDamage = sum(propDamage),
                    cropDamage = sum(cropDamage))
```

Omitting events that did not cause financial damage

```
econDamage <- econDamage[(econDamage$propDamage > 0 | econDamage$cropDamage > 0), ]
```

sorting the data in decreasing order

```
propDmgSorted <- econDamage[order(econDamage$propDamage, decreasing = T), ]
cropDmgSorted <- econDamage[order(econDamage$cropDamage, decreasing = T), ]
```

3. Results

3.1 Effects on population health

Lists of the Top 5 weather events affecting the populations health (injuries and deaths).

```
head(injury[, c("EVTYPE", "injuries")],5)
```

```
##             EVTYPE injuries
## 834        TORNADO    91346
## 856      TSTM WIND     6957
## 170          FLOOD     6789
## 130 EXCESSIVE HEAT     6525
## 464      LIGHTNING     5230
```

```
head(fatal[, c("EVTYPE", "fatalities")],5)
```

```
##             EVTYPE fatalities
```

```
## 834        TORNADO       5633
## 130 EXCESSIVE HEAT       1903
## 153    FLASH FLOOD        978
## 275           HEAT        937
## 464      LIGHTNING        816
```

Plotting the Top 5 population health events:

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.2.5
```

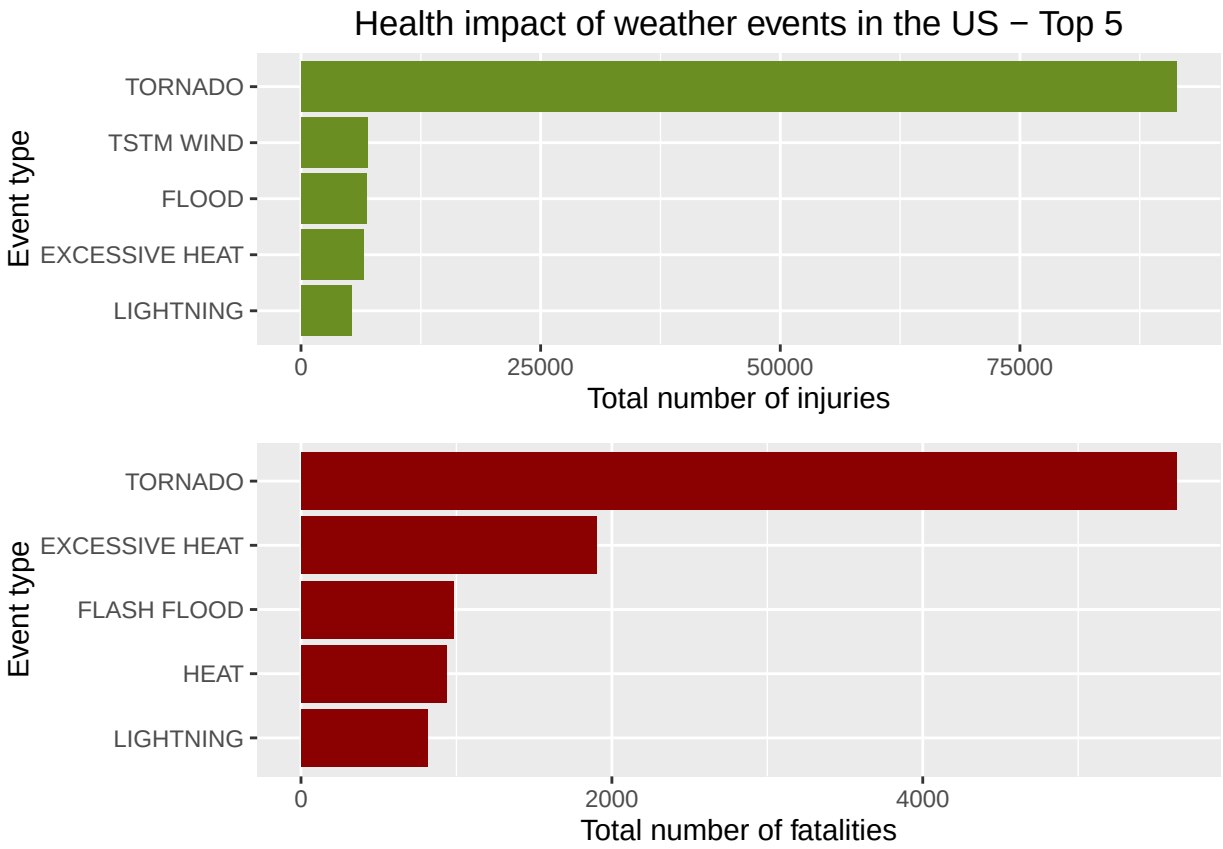
```
library(grid)
library(gridExtra)
```

```
## Warning: package 'gridExtra' was built under R version 3.2.5
```

```
p1 <- ggplot(data=head(injury,5), aes(x=reorder(EVTYPE, injuries), y=injuries)) +
   geom_bar(fill="olivedrab",stat="identity")  + coord_flip() +
    ylab("Total number of injuries") + xlab("Event type") +
    ggtitle("Health impact of weather events in the US - Top 5") +
    theme(legend.position="none")

p2 <- ggplot(data=head(fatal,5), aes(x=reorder(EVTYPE, fatalities), y=fatalities)) +
    geom_bar(fill="red4",stat="identity") + coord_flip() +
    ylab("Total number of fatalities") + xlab("Event type") +
    theme(legend.position="none")

grid.arrange(p1, p2, nrow =2)
```



The plots shows that tornados are by far the most dangerous events when it comes to population health.

3.2. Economic Consequences

What are the top 5 weather events that have the greatest economic consequences?

```
head(propDmgSorted[, c("EVTYPE", "propDamage")], 5)
```

```
##                 EVTYPE   propDamage
## 153        FLASH FLOOD 6.820237e+13
## 786 THUNDERSTORM WINDS 2.086532e+13
## 834            TORNADO 1.078951e+12
## 244               HAIL 3.157558e+11
## 464          LIGHTNING 1.729433e+11
```

```
head(cropDmgSorted[, c("EVTYPE", "cropDamage")], 5)
```

```
##           EVTYPE  cropDamage
## 95       DROUGHT 13972566000
## 170        FLOOD  5661968450
## 590 RIVER FLOOD  5029459000
## 427   ICE STORM  5022113500
## 244         HAIL  3025974480
```

Plotting the Top 5 property damage and crop damage events:

```
p1 <- ggplot(data=head(propDmgSorted,5), aes(x=reorder(EVTYPE, propDamage),
          y=log10(propDamage), fill=propDamage )) +
```

```
    geom_bar(fill="darkred", stat="identity") + coord_flip() +
    xlab("Event type") + ylab("Property damage in dollars (log10)") +
    ggtitle("Economic impact of weather events in the US - Top 5") +
    theme(plot.title = element_text(hjust = 0))

p2 <- ggplot(data=head(cropDmgSorted,5), aes(x=reorder(EVTYPE, cropDamage),
                                             y=cropDamage, fill=cropDamage)) +
    geom_bar(fill="goldenrod", stat="identity") + coord_flip() +
    xlab("Event type") + ylab("Crop damage in dollars") +
    theme(legend.position="none")

grid.arrange(p1, p2, ncol=1, nrow =2)
```