Storm Data Analysis

Rohit Jain

8/28/2018

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.In this data analysis we will be analysing the effect of these events on human and property with help of R.

Data Processing

```
#loading the dataset
dt <- read.table("StormData.csv.bz2", sep = ",", header = TRUE)

#preprocessing the Event type:
dt$EVTYPE <- tolower(dt$EVTYPE)
dt$EVTYPE[grepl(pattern = "*tstm*", dt$EVTYPE)] <- "Tstm Wind"
dt$EVTYPE[grepl(pattern = "^tropical strom*", dt$EVTYPE)] <- "Tropical Storm"
dt$EVTYPE[grepl(pattern = "*flood*", dt$EVTYPE)] <- "Flood"
dt$EVTYPE[grepl(pattern = "*hail*", dt$EVTYPE)] <- "Hail"
dt$EVTYPE[grepl(pattern = "*rain*", dt$EVTYPE)] <- "Rain"
dt$EVTYPE[grepl(pattern = "*thunder*", dt$EVTYPE)] <- "Thunderstorm"
dt$EVTYPE[grepl(pattern = "*snow*", dt$EVTYPE)] <- "Snow Related"
dt$EVTYPE[grepl(pattern = "*surf*", dt$EVTYPE)] <- "Surf Related"
dt$EVTYPE[grepl(pattern = "*microburst*", dt$EVTYPE)] <- "Dry Microburst"
dt$EVTYPE[grepl(pattern = "*funnel*", dt$EVTYPE)] <- "Funnel Cloud"
dt$EVTYPE[grepl(pattern = "*glaze*", dt$EVTYPE)] <- "Glaze"
dt$EVTYPE[grepl(pattern = "*shower*", dt$EVTYPE)] <- "Rain"
dt$EVTYPE[grepl(pattern = "*high wind*", dt$EVTYPE)] <- "High Winds"
dt$EVTYPE[grepl(pattern = "*hurricane*", dt$EVTYPE)] <- "Hurricane"
dt$EVTYPE[grepl(pattern = "*light*", dt$EVTYPE)] <- "Lighting"
dt$EVTYPE[grepl(pattern = "*summary*", dt$EVTYPE)] <- "Summary related Entry"
dt$EVTYPE[grepl(pattern = "*tornado*", dt$EVTYPE)] <- "Tornado"
dt$EVTYPE[grepl(pattern = "*waterspout*", dt$EVTYPE)] <- "Waterspout"
dt$EVTYPE[grepl(pattern = "*wild*", dt$EVTYPE)] <- "Wild Fire"
dt$EVTYPE[grepl(pattern = "*record*", dt$EVTYPE)] <- "Record Conditions"
dt$EVTYPE[grepl(pattern = "*blizzard*", dt$EVTYPE)] <- "Blizzard Conditions"
dt$EVTYPE[grepl(pattern = "*cold*", dt$EVTYPE)] <- "Cold Related"
dt$EVTYPE[grepl(pattern = "*extreme*", dt$EVTYPE)] <- "Extreme Conditions"
dt$EVTYPE[grepl(pattern = "*frost*", dt$EVTYPE)] <- "Frost"
dt$EVTYPE[grepl(pattern = "*gusty*", dt$EVTYPE)] <- "Gusty Wind"
dt$EVTYPE[grepl(pattern = "^heat*", dt$EVTYPE)] <- "Heat Related"

# Pre processing for economic effect:

dt$PROPDMGEXP <- as.character(dt$PROPDMGEXP)
dt$PROPDMGEXP[dt$PROPDMGEXP %in% c("m","M")] <- 6
```

```
dt$PROPDMGEXP[dt$PROPDMGEXP %in% c("k","K")] <- 3
dt$PROPDMGEXP[dt$PROPDMGEXP %in% c("h","H")] <- 2
dt$PROPDMGEXP[dt$PROPDMGEXP %in% c("b","B")] <- 9
dt$PROPDMGEXP[dt$PROPDMGEXP %in% c('-','?','+',' ')] <- 0
dt$PROPDMGEXP <- as.numeric(dt$PROPDMGEXP)
dt$propdmg <- dt$PROPDMG * 10^dt$PROPDMGEXP

dt$CROPDMGEXP <- as.character(dt$CROPDMGEXP)
dt$CROPDMGEXP[dt$CROPDMGEXP %in% c("m","M")] <- 6
dt$CROPDMGEXP[dt$CROPDMGEXP %in% c("k","K")] <- 3
dt$CROPDMGEXP[dt$CROPDMGEXP %in% c("h","H")] <- 2
dt$CROPDMGEXP[dt$CROPDMGEXP %in% c("b","B")] <- 9
dt$CROPDMGEXP[dt$CROPDMGEXP %in% c('-','?','+',' ')] <- 0
dt$CROPDMGEXP <- as.numeric(dt$CROPDMGEXP)
dt$cropdmg <- dt$CROPDMG * 10^dt$CROPDMGEXP
```

Data Analysis

Effect of the Events on Humans(Fatalities and Inturies)

In this part of the document we will be extracting out the top 10 event for the Dataset, which has the worst effect on human in context of fatality and Injury.

```
library(plyr)
humanEffect <- ddply(dt, .(EVTYPE), summarize,
                     fatalities = sum(FATALITIES),
                     injuries= sum(INJURIES))
#Event type
fatal_events <- head(humanEffect[order(humanEffect$fatalities, decreasing = T), ], 10)
injury_events <- head(humanEffect[order(humanEffect$injuries, decreasing = T), ], 10)
```

Top 10 Events having worst effect on human health(Injuries and Fatalities)

```
fatal_events[, c("EVTYPE", "fatalities")]
```

```
##             EVTYPE fatalities
## 172        Tornado       5636
## 55  excessive heat       1903
## 61           Flood       1525
## 78    Heat Related       1121
## 111       Lighting        817
## 180      Tstm Wind        544
## 27    Cold Related        435
## 146    rip current        368
## 87      High Winds        294
## 11       avalanche        224
```

```
injury_events[, c("EVTYPE", "injuries")]
```

```
##             EVTYPE injuries
## 172        Tornado    91407
## 61           Flood     8604
## 180      Tstm Wind     7065
```

```
## 55  excessive heat     6525
## 111       Lighting     5231
## 78    Heat Related     2494
## 170   Thunderstorm     2479
## 102      ice storm     1975
## 224      Wild Fire     1606
## 87      High Winds     1480
```

Effect of the Events on Property and Crops

In this part of the document we will be extracting out the top 10 event for the Dataset, which has the worst economic effect on Property and Crops.

```
ecoEffect <- ddply(dt, .(EVTYPE), summarize,
                   property = sum(propdmg),
                   crops= sum(cropdmg))

prop_dmg <- head(ecoEffect[order(ecoEffect$property, decreasing = TRUE),],10)
prop_dmg$property[is.na(prop_dmg$property)] <- 0
crop_dmg <- head(ecoEffect[order(ecoEffect$crops, decreasing = TRUE ),],10)
crop_dmg$crops[is.na(crop_dmg$crops)] <- 0
```

Top 10 Events having worst economic effect

```
prop_dmg[, c("EVTYPE", "property")]
```

```
##                     EVTYPE property
## 63            forest fires  5000000
## 176 tropical storm alberto  5000000
## 103       ice/strong winds  3500000
## 227          wind and wave  1000000
## 24         coastal erosion   766000
## 157        sleet/ice storm   500000
## 178  tropical storm gordon   500000
## 117     marine strong wind   418330
## 9    astronomical low tide   320000
## 230             wind storm   300000
```

```
crop_dmg[, c("EVTYPE", "crops")]
```

```
##                     EVTYPE    crops
## 58      excessive wetness 1.42e+08
## 28           cool and wet 5.00e+06
## 63           forest fires 5.00e+05
## 178 tropical storm gordon 5.00e+05
## 9   astronomical low tide 0.00e+00
## 210      volcanic ashfall 0.00e+00
## 1                    wind 0.00e+00
## 2                       ? 0.00e+00
## 3         abnormal warmth 0.00e+00
## 4          abnormally dry 0.00e+00
```

Result

In this section we will be depicting the outcome of the analysis in the form graphs where the "Event Vs Effect" relationship is shown.

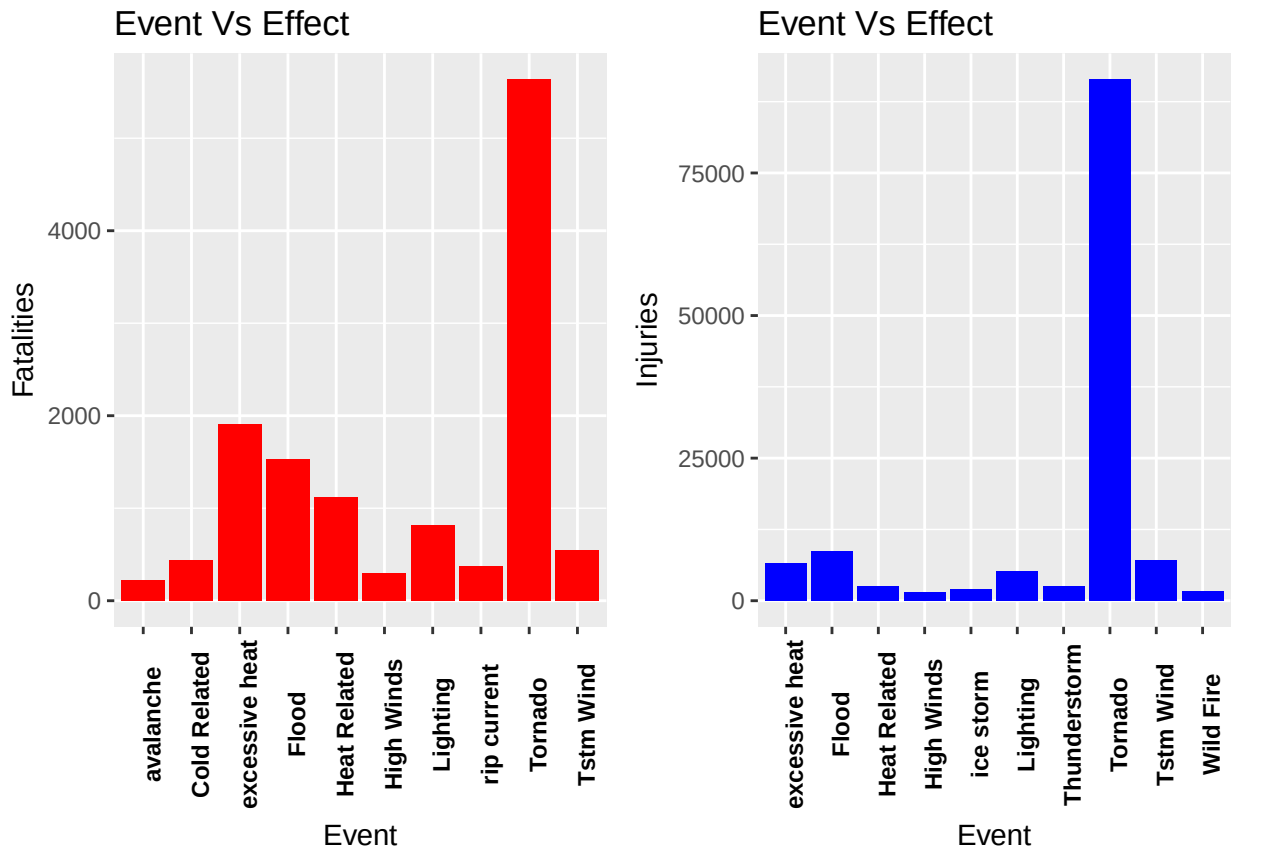
Human Effect

```
library(ggplot2)
library(gridExtra)

fatalPlot <- ggplot(data = fatal_events, aes(EVTYPE, fatalities)) +
             geom_bar(stat = "identity", fill = "red") +
             labs(title = "Event Vs Effect", x = "Event" , y = "Fatalities") +
             theme(axis.text.x = element_text(angle = 90, face = "bold", colour = "black"))

injuryPlot <- ggplot(data = injury_events, aes(EVTYPE, injuries)) +
             geom_bar(stat = "identity", fill = "blue") +
             labs(title = "Event Vs Effect", x = "Event" , y = "Injuries") +
             theme(axis.text.x = element_text(angle = 90, face = "bold", colour = "black"))

grid.arrange(fatalPlot, injuryPlot, nrow = 1)
```



Economic Effect

```
library(ggplot2)
library(gridExtra)

propPlot <- ggplot(data = prop_dmg, aes(EVTYPE, log(property))) +
  geom_bar(stat = "identity", fill = "red") +
  labs(title = "Event Vs Effect", x = "Event" , y = "Property damage") +
  theme(axis.text.x = element_text(angle = 90, face = "bold", colour = "black"))

cropPlot <- ggplot(data = crop_dmg, aes(EVTYPE, log(crops))) +
  geom_bar(stat = "identity", fill = "blue") +
  labs(title = "Event Vs Effect", x = "Event" , y = "Crop damage") +
  theme(axis.text.x = element_text(angle = 90, face = "bold", colour = "black"))

grid.arrange(propPlot, cropPlot, nrow = 1)
```