Reproducible Research: Peer Assessment 2

Analysis Severe Weather Impact on Public Health and Economy

Synonpsis

Based on the storm data collected from the U.S. National Oceanic and Atmospheric Administration's (NOAA) from 1950 - 2011, this report analyzes the impact of different weather events on public health and economy. It is to answer to two questions:
1. Across the United States, which types of events (as indicated in the EVTYPE variable) are most harmful with respect to population health?
2. Across the United States, which types of events have the greatest economic consequences?

Data Processing

```
library(R.utils)
```

```
## Loading required package: R.oo
## Loading required package: R.methodsS3
## R.methodsS3 v1.6.1 (2014-01-04) successfully loaded. See ?R.methodsS3 for help.
## R.oo v1.18.0 (2014-02-22) successfully loaded. See ?R.oo for help.
##
## Attaching package: 'R.oo'
##
## The following objects are masked from 'package:methods':
##
##     getClasses, getMethods
##
## The following objects are masked from 'package:base':
##
##     attach, detach, gc, load, save
##
## R.utils v1.32.4 (2014-05-14) successfully loaded. See ?R.utils for help.
##
## Attaching package: 'R.utils'
##
## The following object is masked from 'package:utils':
##
##     timestamp
##
## The following objects are masked from 'package:base':
##
##     cat, commandArgs, getOption, inherits, isOpen, parse, warnings
```

```
library(ggplot2)
library(plyr)
require(gridExtra)
```

```
## Loading required package: gridExtra
## Loading required package: grid
```

```
stormData <- read.csv("stormData.csv")
head(stormData, n=3)
```

Loading the data

```
##   STATE__           BGN_DATE BGN_TIME TIME_ZONE COUNTY COUNTYNAME STATE
## 1       1 4/18/1950 0:00:00     0130       CST     97     MOBILE    AL
## 2       1 4/18/1950 0:00:00     0145       CST      3    BALDWIN    AL
## 3       1 2/20/1951 0:00:00     1600       CST     57    FAYETTE    AL
##    EVTYPE BGN_RANGE BGN_AZI BGN_LOCATI END_DATE END_TIME COUNTY_END
## 1 TORNADO         0                                               0
## 2 TORNADO         0                                               0
## 3 TORNADO         0                                               0
##   COUNTYENDN END_RANGE END_AZI END_LOCATI LENGTH WIDTH F MAG FATALITIES
## 1         NA         0                      14.0   100 3   0          0
## 2         NA         0                       2.0   150 2   0          0
## 3         NA         0                       0.1   123 2   0          0
##   INJURIES PROPDMG PROPDMGEXP CROPDMG CROPDMGEXP WFO STATEOFFIC ZONENAMES
## 1       15    25.0          K       0
## 2        0     2.5          K       0
## 3        2    25.0          K       0
##   LATITUDE LONGITUDE LATITUDE_E LONGITUDE_ REMARKS REFNUM
## 1     3040      8812       3051       8806              1
## 2     3042      8755          0          0              2
## 3     3340      8742          0          0              3
```

Preparing fatalities and injuries data

```
sortAid <- function(fieldName, top = 15, dataset = stormData) {
    index <- which(colnames(dataset) == fieldName)
    field <- aggregate(dataset[, index], by = list(dataset$EVTYPE), FUN = "sum")
    names(field) <- c("EVTYPE", fieldName)
    field <- arrange(field, field[, 2], decreasing = T)
    field <- head(field, n = top)
    field <- within(field, EVTYPE <- factor(x = EVTYPE, levels = field$EVTYPE))
    return(field)
}

fatalities <- sortAid("FATALITIES", dataset = stormData)
injuries <- sortAid("INJURIES", dataset = stormData)
```

Preparing property and crop damage data

```
convertAid <- function(dataset = stormData, fieldName, newFieldName) {
    totalLen <- dim(dataset)[2]
    index <- which(colnames(dataset) == fieldName)
    dataset[, index] <- as.character(dataset[, index])
    logic <- !is.na(toupper(dataset[, index]))
    dataset[logic & toupper(dataset[, index]) == "B", index] <- "9"
    dataset[logic & toupper(dataset[, index]) == "M", index] <- "6"
    dataset[logic & toupper(dataset[, index]) == "K", index] <- "3"
```

```
    dataset[logic & toupper(dataset[, index]) == "H", index] <- "2"
    dataset[logic & toupper(dataset[, index]) == "", index] <- "0"
    dataset[, index] <- as.numeric(dataset[, index])
    dataset[is.na(dataset[, index]), index] <- 0
    dataset <- cbind(dataset, dataset[, index - 1] * 10^dataset[, index])
    names(dataset)[totalLen + 1] <- newFieldName
    return(dataset)
}

stormData <- convertAid(stormData, "PROPDMGEXP", "propertyDamage")
```

```
## Warning: NAs introduced by coercion
```

```
stormData <- convertAid(stormData, "CROPDMGEXP", "cropDamage")
```

```
## Warning: NAs introduced by coercion
```

```
names(stormData)
```

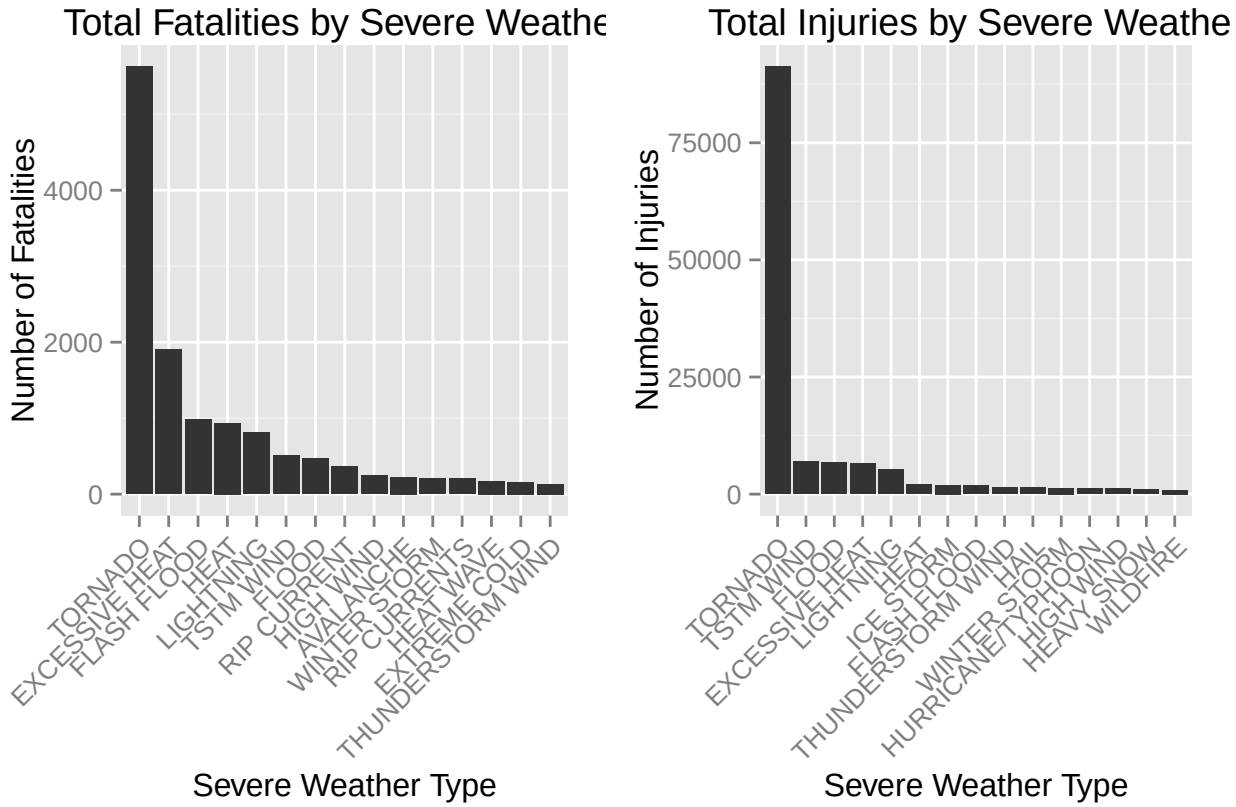
```
##  [1] "STATE__"        "BGN_DATE"       "BGN_TIME"       "TIME_ZONE"
##  [5] "COUNTY"         "COUNTYNAME"     "STATE"          "EVTYPE"
##  [9] "BGN_RANGE"      "BGN_AZI"        "BGN_LOCATI"     "END_DATE"
## [13] "END_TIME"       "COUNTY_END"     "COUNTYENDN"     "END_RANGE"
## [17] "END_AZI"        "END_LOCATI"     "LENGTH"         "WIDTH"
## [21] "F"              "MAG"            "FATALITIES"     "INJURIES"
## [25] "PROPDMG"        "PROPDMGEXP"     "CROPDMG"        "CROPDMGEXP"
## [29] "WFO"            "STATEOFFIC"     "ZONENAMES"      "LATITUDE"
## [33] "LONGITUDE"      "LATITUDE_E"     "LONGITUDE_"     "REMARKS"
## [37] "REFNUM"         "propertyDamage" "cropDamage"
```

```
property <- sortAid("propertyDamage", dataset = stormData)
crop <- sortAid("cropDamage", dataset = stormData)
property$propertyDamage <- property$propertyDamage/10^9
crop$cropDamage <- crop$cropDamage/10^9
```

Results

```
fatalitiesPlot <- qplot(EVTYPE, data = fatalities, weight = FATALITIES, geom = "bar", binwidth = 1) +
    scale_y_continuous("Number of Fatalities") +
    theme(axis.text.x = element_text(angle = 45,
    hjust = 1)) + xlab("Severe Weather Type") +
    ggtitle("Total Fatalities by Severe Weather")
injuriesPlot <- qplot(EVTYPE, data = injuries, weight = INJURIES, geom = "bar", binwidth = 1) +
    scale_y_continuous("Number of Injuries") +
    theme(axis.text.x = element_text(angle = 45,
    hjust = 1)) + xlab("Severe Weather Type") +
    ggtitle("Total Injuries by Severe Weather")
grid.arrange(fatalitiesPlot, injuriesPlot, ncol = 2)
```

Aross the United States, Which types of events are nost harmful with respect to population



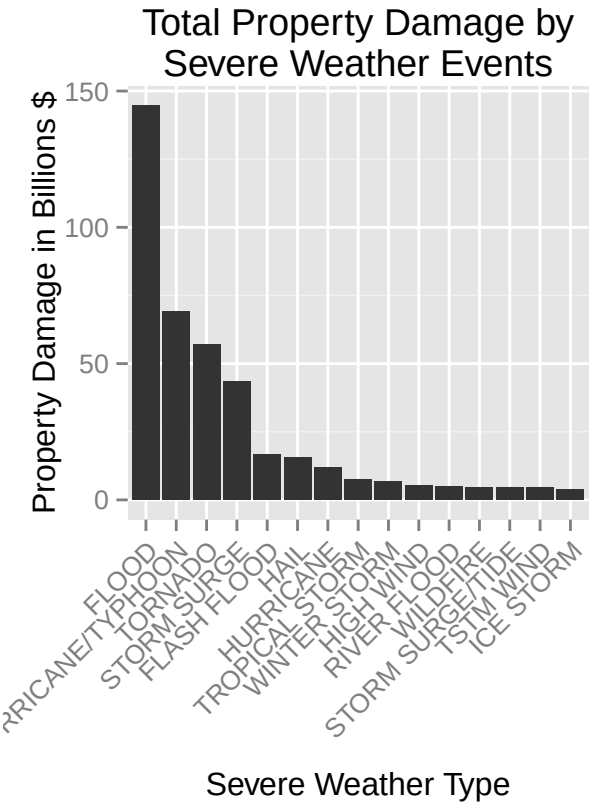
health?

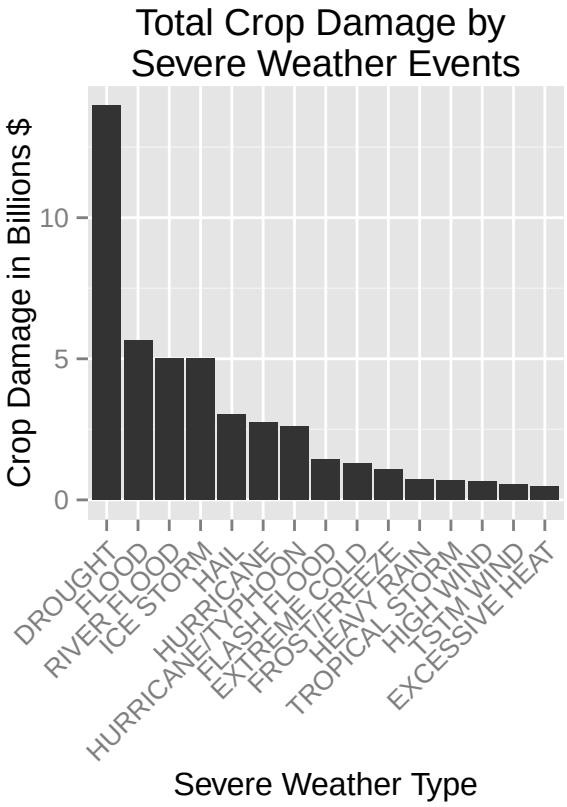
The most harmful weather event to population health is Tornado.It is cause for both the highest fatalities and the highest injuries across United States.

```
propertyPlot <- qplot(EVTYPE, data = property, weight = propertyDamage, geom = "bar", binwidth = 1) +
    theme(axis.text.x = element_text(angle = 45, hjust = 1)) + scale_y_continuous("Property Damage in Bi
    xlab("Severe Weather Type") + ggtitle("Total Property Damage by\n Severe Weather Events ")

cropPlot<- qplot(EVTYPE, data = crop, weight = cropDamage, geom = "bar", binwidth = 1) +
    theme(axis.text.x = element_text(angle = 45, hjust = 1)) + scale_y_continuous("Crop Damage in Billi
    xlab("Severe Weather Type") + ggtitle("Total Crop Damage by \nSevere Weather Events")
grid.arrange(propertyPlot, cropPlot, ncol = 2)
```

Across the United States, which types of events have the greatest economic consequences?





The weather events have the greatest economic consequences are: flood, drought, Tornado and Typhoon.

Across the United States, flood, tornado and typhoon have caused the greatest damage to properties. Drought and flood come as the causes for the greatest damage to crops.