

# Reproducible Research: Peer Assessment 2

*Vince Glick*

*Sunday, February 22, 2015*

## Synopsis

The basic goal of this assignment is to explore the NOAA Storm Database and answer some basic questions about severe weather events. The following questions will be answered: 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

Load/Import Required Libraries & Data Files

```
echo = TRUE
options(scipen = 1)
require(ggplot2)
```

```
## Loading required package: ggplot2
```

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

Aggregate and clean the data for Fatalities and Injuries nationwide.

```
health_Aggregate<-aggregate(x=storm_Data[,c("FATALITIES", "INJURIES")],
                             by=list(storm_Data$EVTYPE), FUN=sum)
health_Aggregate$Total.Harmful.Health<-health_Aggregate$FATALITIES+health_Aggregate$INJURIE
S
health_Clean<-health_Aggregate[health_Aggregate$Total.Harmful.Health!=0, ]

Total.Harmful.Health<-health_Clean$Total.Harmful.Health

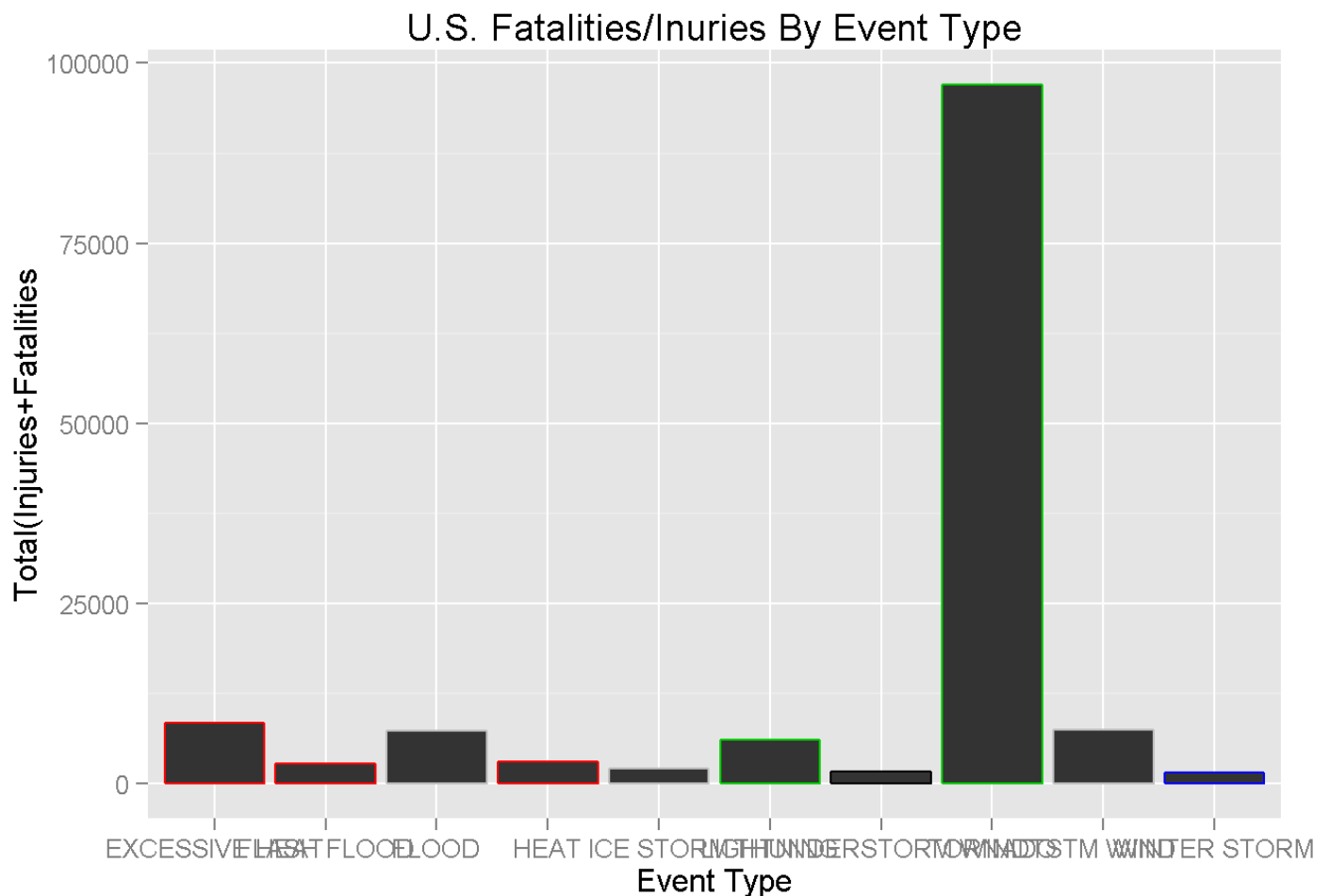
health_Clean_Order<-health_Clean[order(-Total.Harmful.Health),]
health_Clean_Order<-health_Clean_Order[1:10,]
```

Assign Variables for ease of reference for the remainder of this portion

```
fatalities<-health_Clean_Order$FATALITIES
injuries<-health_Clean_Order$INJURIES
evtype<-health_Clean_Order$Group.1
```

Plot the graph using ggplot for total Fatalities & Injuries By Event Type. Here, a bar plot is used to display the top 10 most event types most detrimental to injuries & fatalities.

```
ggplot(data=health_Clean_Order, aes(x=evtype, y= Total.Harmful.Health,)) +
  geom_bar(stat="identity", colour=evtype) +
  xlab("Event Type") + ylab("Total(Injuries+Fatalities)") +
  ggtitle("U.S. Fatalities/Inuries By Event Type")
```



## Process the data for Economic Impacts Nationwide

Based on the data from the “NATIONAL WEATHER SERVICE INSTRUCTION 10-1605” partnered with this assignment, the values provided for Property Damage and Crop Damage will be converted as required. (This publication is available at <http://www.nws.noaa.gov/directives/> (<http://www.nws.noaa.gov/directives/>))

```

storm_Data$PROPDMGEXP = as.character(storm_Data$PROPDMGEXP)
storm_Data$PROPDMGEXP[toupper(storm_Data$PROPDMGEXP) == 'B'] = "9"
storm_Data$PROPDMGEXP[toupper(storm_Data$PROPDMGEXP) == 'M'] = "6"
storm_Data$PROPDMGEXP[toupper(storm_Data$PROPDMGEXP) == 'K'] = "3"
storm_Data$PROPDMGEXP[toupper(storm_Data$PROPDMGEXP) == 'H'] = "2"
storm_Data$PROPDMGEXP = as.numeric(storm_Data$PROPDMGEXP)

```

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

```

storm_Data$PROPDMGEXP[is.na(storm_Data$PROPDMGEXP)] = 0
storm_Data$PropertyDamage = storm_Data$PROPDMG * 10^storm_Data$PROPDMGEXP
summary(storm_Data$PropertyDamage)

```

```

##      Min.   1st Qu.   Median     Mean   3rd Qu.    Max.
## 0.000e+00 0.000e+00 0.000e+00 4.746e+05 5.000e+02 1.150e+11

```

```

storm_Data$CROPDMGEXP = as.character(storm_Data$CROPDMGEXP)
storm_Data$CROPDMGEXP[toupper(storm_Data$CROPDMGEXP) == 'B'] = "9"
storm_Data$CROPDMGEXP[toupper(storm_Data$CROPDMGEXP) == 'M'] = "6"
storm_Data$CROPDMGEXP[toupper(storm_Data$CROPDMGEXP) == 'K'] = "3"
storm_Data$CROPDMGEXP[toupper(storm_Data$CROPDMGEXP) == 'H'] = "2"
storm_Data$CROPDMGEXP[toupper(storm_Data$CROPDMGEXP) == ''] = "0"
storm_Data$CROPDMGEXP = as.numeric(storm_Data$CROPDMGEXP)

```

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

```

storm_Data$CROPDMGEXP[is.na(storm_Data$CROPDMGEXP)] = 0
storm_Data$CropDamage = storm_Data$CROPDMG * 10^storm_Data$CROPDMGEXP
summary(storm_Data$CropDamage)

```

```

##      Min.   1st Qu.   Median     Mean   3rd Qu.    Max.
##      0         0         0    54420         0 5000000000

```

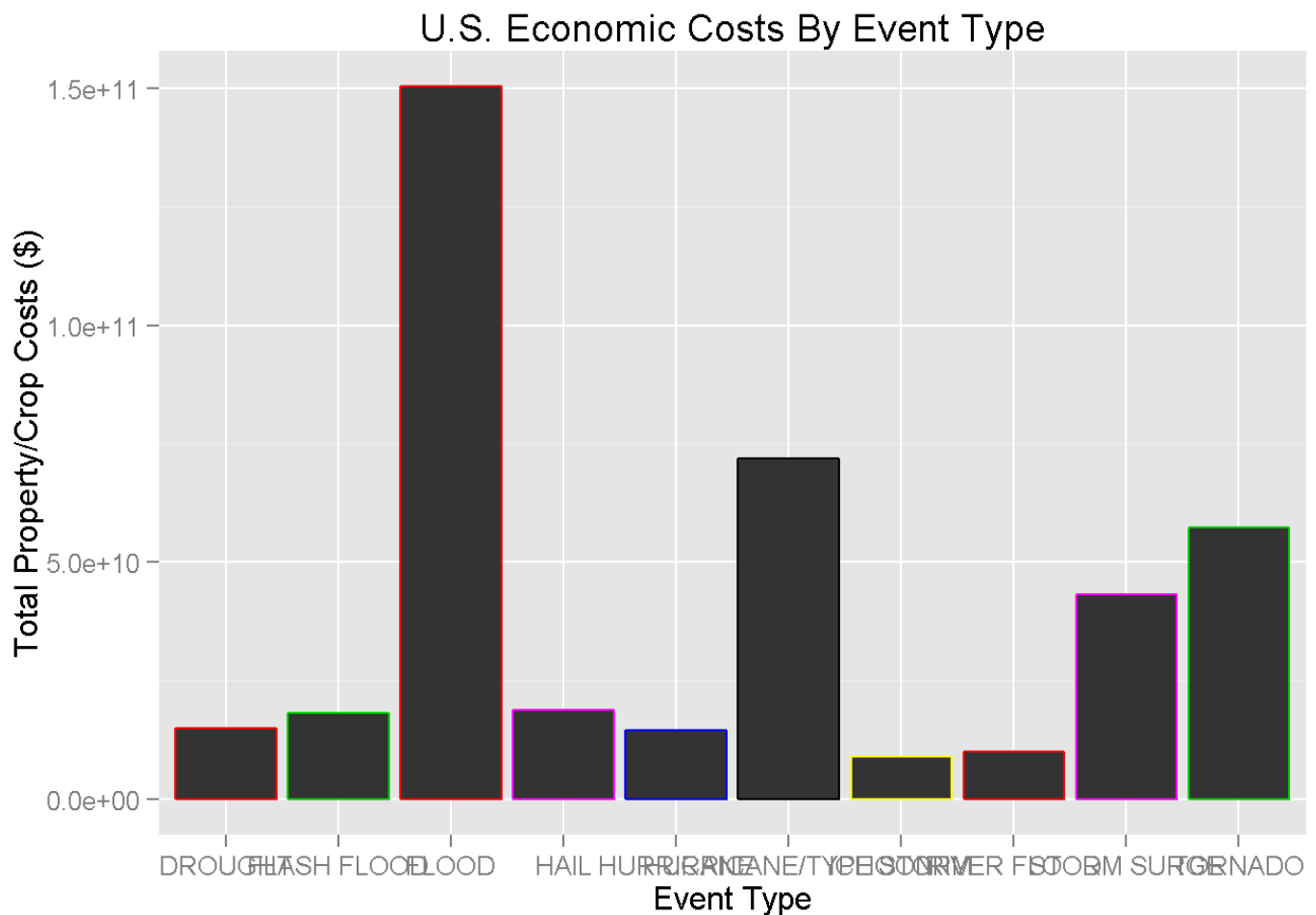
```

damage_Aggregate<-aggregate(x=storm_Data[,c("PropertyDamage","CropDamage")],
                             by=list(storm_Data$EVTYPE), FUN=sum)
damage_Aggregate$total_Damage<-damage_Aggregate$PropertyDamage + damage_Aggregate$CropDamage
damage_Aggregate<-damage_Aggregate[order(-damage_Aggregate$total_Damage),]
damage_Aggregate<-damage_Aggregate[1:10,]

```

Plot the graph using ggplot for the total Economic Impacts By Event Type. A bar plot is applied to display the top 10 Event types that had the largest economic impacts.

```
ggplot(data=damage_Aggregate, aes(x=damage_Aggregate$Group.1, y=damage_Aggregate$total_Damage,)) + geom_bar(stat="identity", colour=damage_Aggregate$Group.1) +  
  xlab("Event Type") + ylab("Total Property/Crop Costs ($)") +  
  ggtitle("U.S. Economic Costs By Event Type")
```



## Results

Our findings suggest that the largest impacts to injuries and fatalities as a whole across the nation are Tornadoes and Excessive Heat. Furthermore, based on the data, Floods and Hurricanes/Typhoons have the largest economic costs as a total of property and crops nationwide.