

# Title: “Reproducible Research - Project 2 - Weather Event Analysis”

## Synopsis:

1. In order to analyze the impact of weather on people and property, we are using data from the National Weather Service
2. The data was summarized according to injuries, fatalities, property and crop damage reported as the result of specific weather events.
3. The data was further summarized to reveal the top weather events in the categories mentioned above.
4. Some level of data cleaning was performed to assure the damage dollar amounts were expressed in millions.
5. Histogram diagrams were used to reveal the storm events that were highest in each category.

## Conclusions:

1. Tornadoes appear to be the leader in 3 of the 4 categories with the exception of Crop Damage which Droughts are the main cause of damage.
2. Other relevant events that are close behind Tornadoes Heat/Excessive Heat are responsible for large amount of fatalities, injuries and crop damage.
3. Floods/Flash Floods were second highest in property damage.

output: pdf\_document: default html\_document: default word\_document: default

## Raw Data:

Storm Data Location - <https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2>

Data Description - National Weather Service Storm Data Documentation - [https://d396qusza40orc.cloudfront.net/repdata%2Fpeer2\\_\\_doc%2Fpd01016005curr.pdf](https://d396qusza40orc.cloudfront.net/repdata%2Fpeer2__doc%2Fpd01016005curr.pdf)

National Climatic Data Center Storm Events FAQ- [https://d396qusza40orc.cloudfront.net/repdata%2Fpeer2\\_\\_doc%2FNCDCC%20Storm%20Events-FAQ%20Page.pdf](https://d396qusza40orc.cloudfront.net/repdata%2Fpeer2__doc%2FNCDCC%20Storm%20Events-FAQ%20Page.pdf)

## Data Processing:

**Summarize the Event Types and the Fatalities, Injuries, Delayed Fatalities, Indirect Fatalities/Injuries and Property Damage, Crop Damage, Other Costs, Delayed Damage, Indirect Damage**

```
knitr::opts_chunk$set(echo = TRUE)

ag<-aggregate(cbind(INJURIES, FATALITIES, PROPDMG, CROPDMG) ~ EVTYPE, data = storm_info, sum)

## Including Plots
#Create Table without the Events that have no fatalities, Injuries
#Limit the Plots to the largest numbers
```

```
plt_fatal<- subset(ag, FATALITIES != 0, select= c(EVTYPE,FATALITIES))
plt_inj<- subset(ag, INJURIES != 0, select= c(EVTYPE,INJURIES))
plt_cropdmg<-subset(ag, CROPDMG != 0, select= c(EVTYPE,CROPDMG))
plt_propdmg<-subset(ag, PROPDMG != 0, select= c(EVTYPE,PROPDMG))

#now sort so we can just get the top 10 items for the grap
plt_fatal<- plt_fatal[ order(plt_fatal$FATALITIES,decreasing = TRUE),]
plt_inj<- plt_inj[ order(plt_inj$INJURIES,decreasing = TRUE),]
plt_cropdmg<- plt_cropdmg[ order(plt_cropdmg$CROPDMG ,decreasing = TRUE),]
plt_propdmg<- plt_propdmg[ order(plt_propdmg$PROPDMG,decreasing = TRUE),]
```

## Results

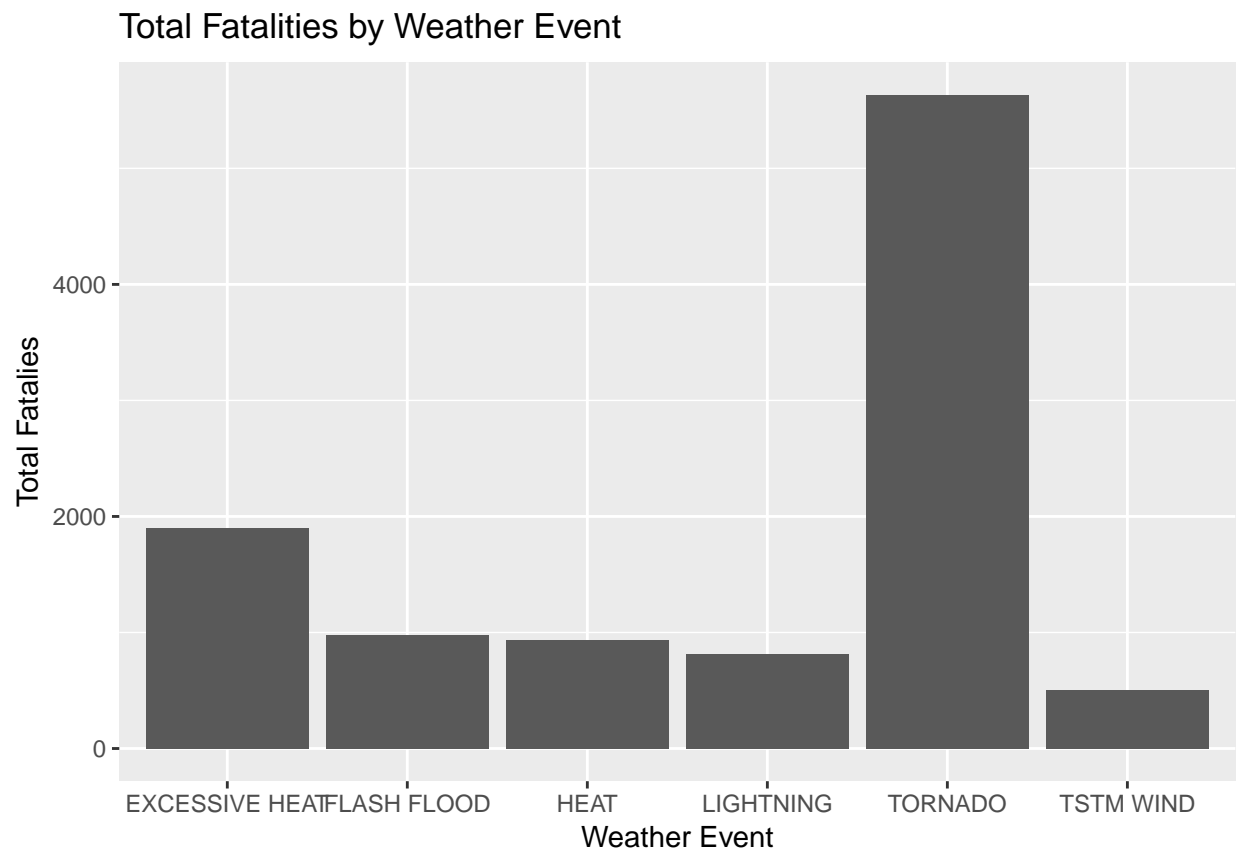
Histogram for Total Fatalities and Injuries by Storm Type

```
require(ggplot2)
```

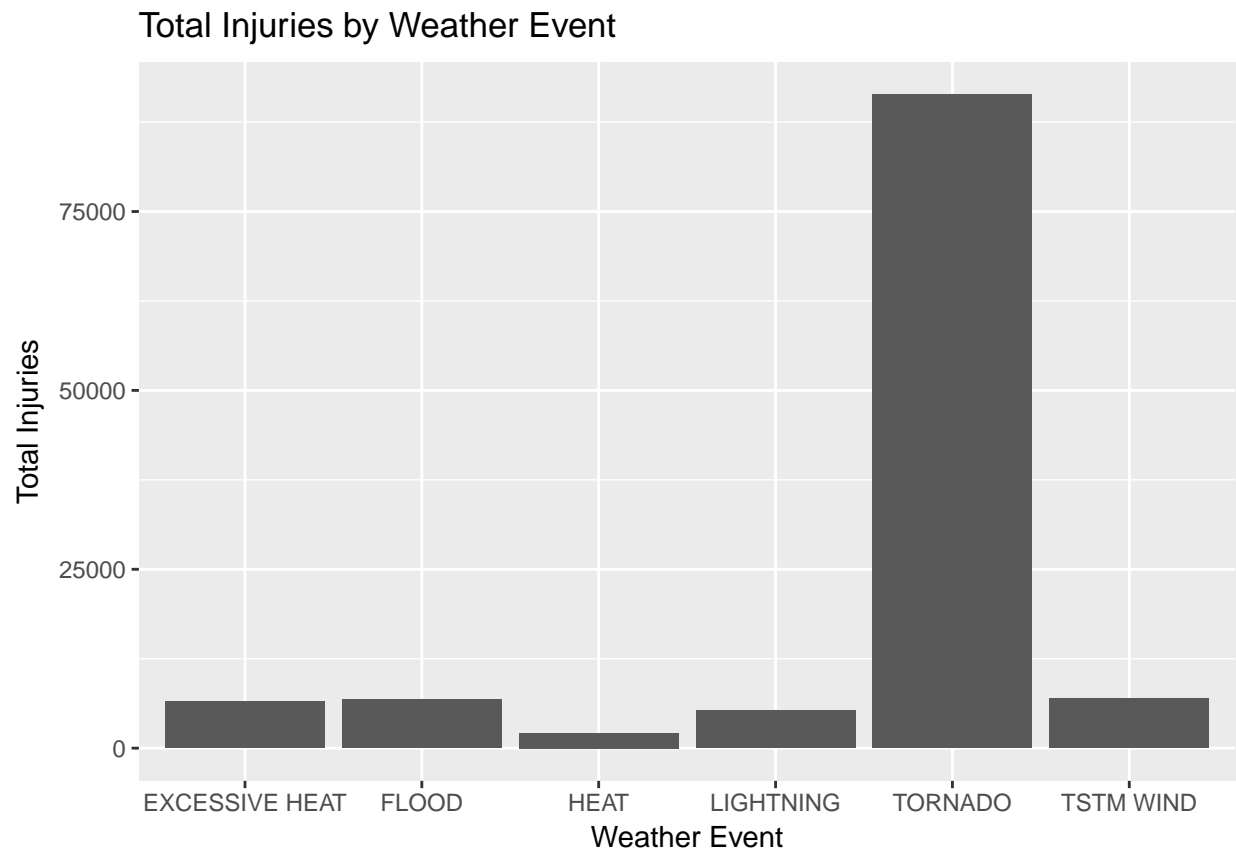
```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 3.3.3
```

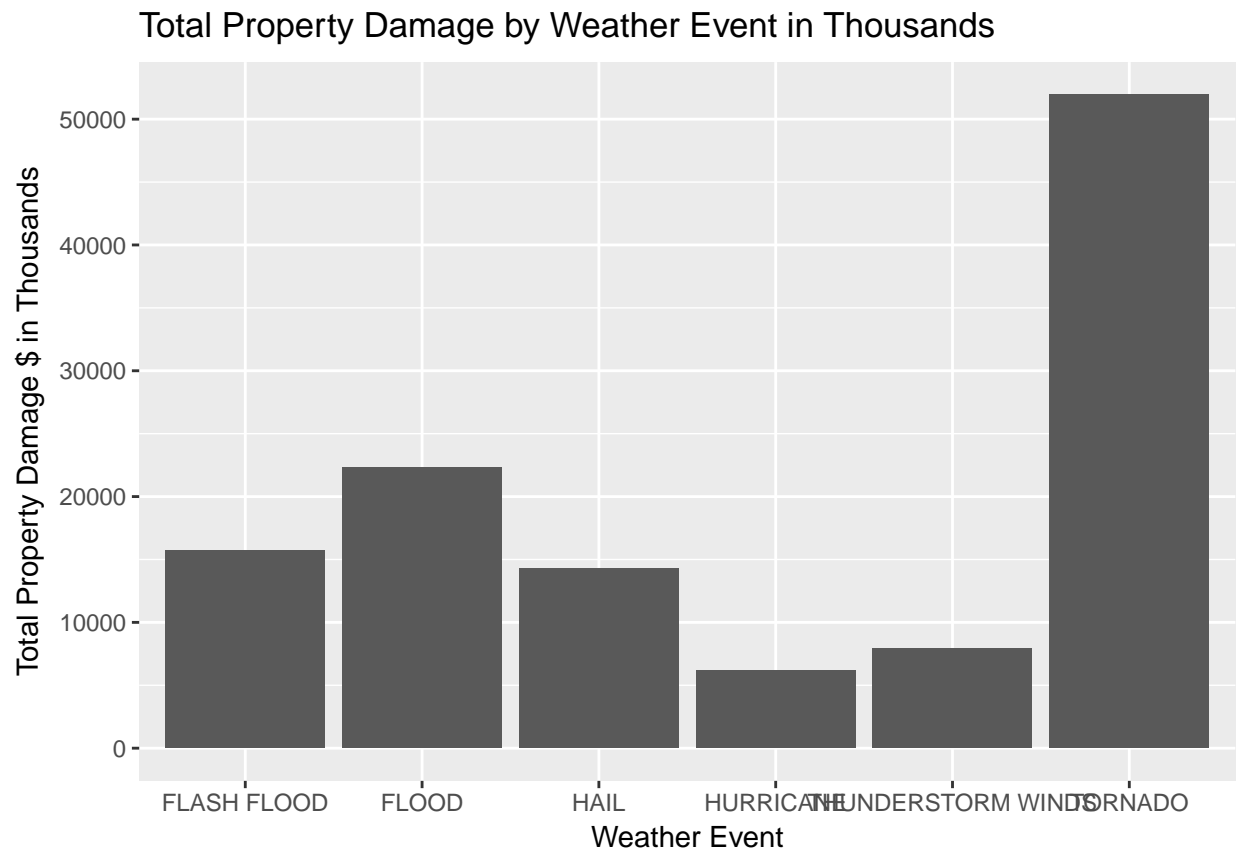
```
ggplot(data=head(plt_fatal), aes(x=EVTYPE,y=FATALITIES))+geom_bar(stat = "identity")+labs(title = "Total Inj
```



```
ggplot(data=head(plt_inj), aes(x=EVTYPE,y=INJURIES))+geom_bar(stat = "identity")+labs(title = "Total Inj
```



```
#Histogram for Total Property and Crop Damage by Storm Type  
ggplot(data=head(plt_propdmg), aes(x=EVTYPE,y=PROPDMG))+geom_bar(stat = "identity")+labs(title = "Total
```



```
ggplot(data=head(plt_cropdmg), aes(x=EVTTYPE,y=CROPDMG))+geom_bar(stat = "identity")+labs(title = "Total
```

Total Crop Damage by Weather Event – In Thousands

