Study on US Weather Events that Impact Health and Economy

Data Processing

The raw data can be found on github or https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2

The code below shows the data processing.

```
storm_df <- read.csv("repdata_data_StormData.csv.bz2")
```

Getting the data for population health. I combined the fatalities and injuries by multiplying the fatality by 10 and then adding it to the number of injuries, and then averaging it over each weather event.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
health_df <- select(storm_df, EVTYPE, FATALITIES, INJURIES)
health_df <- health_df %>% mutate(COMBINED=10*FATALITIES + INJURIES)
health_res <- aggregate(health_df$COMBINED,by=list(Event=health_df$EVTYPE), FUN="mean")
```

Getting the data for economic impact.

```
econ_df <- select(storm_df, EVTYPE, PROPDMG, PROPDMGEXP, CROPDMG, CROPDMGEXP)
# Scale based on the PROPDMGEXP
econ_df$PROPDMGEXP[econ_df$PROPDMGEXP == "K"] <- 1000
econ_df$PROPDMGEXP[econ_df$PROPDMGEXP == "M"] <- 1000000
econ_df$PROPDMGEXP[econ_df$PROPDMGEXP == "B"] <- 1000000000
econ_df$PROPDMGEXP[econ_df$PROPDMGEXP == ""] <- 0
econ_df$CROPDMGEXP[econ_df$CROPDMGEXP == "K"] <- 1000
econ_df$CROPDMGEXP[econ_df$CROPDMGEXP == "M"] <- 1000000
econ_df$CROPDMGEXP[econ_df$CROPDMGEXP == "B"] <- 1000000000
econ_df$CROPDMGEXP[econ_df$CROPDMGEXP == ""] <- 0

econ_df$PROPDMGEXP <- as.numeric(econ_df$PROPDMGEXP)
```

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

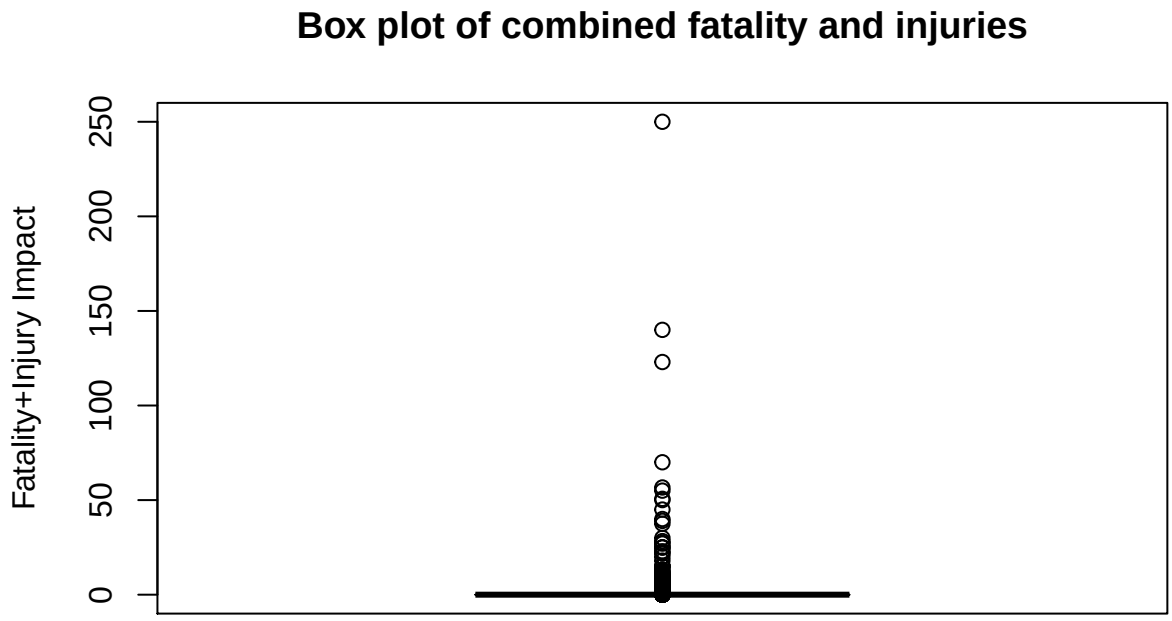
```
econ_df$CROPDMGEXP <- as.numeric(econ_df$CROPDMGEXP)
```

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

```
econ_df <- econ_df %>% mutate(COST= PROPDMG*PROPDMGEXP + CROPDMG*CROPDMGEXP)
econ_res <- aggregate(econ_df$COST,by=list(Event=econ_df$EVTYPE), FUN="mean")
```

Results

```
boxplot(health_res$x, main="Box plot of combined fatality and injuries",
        ylab="Fatality+Injury Impact")
```

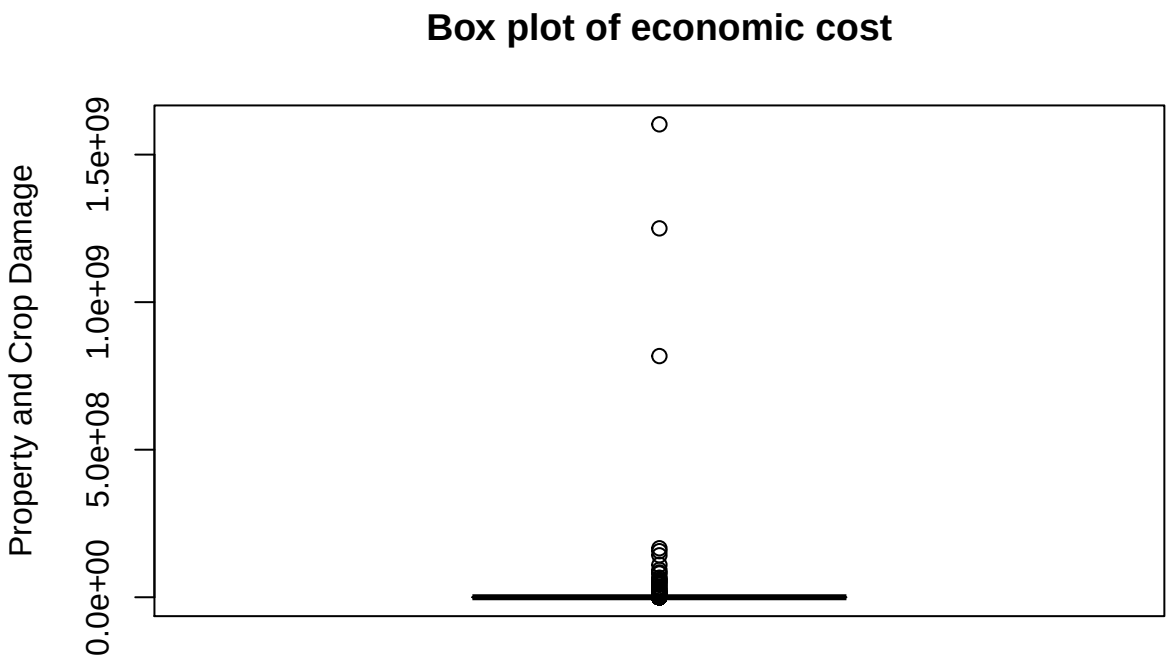


The weather event that is most harmful to the population health is:

```
print(health_res[which.max(health_res$x),1])
```

```
## [1] "TORNADOES, TSTM WIND, HAIL"
```

```
boxplot(econ_res$x, main="Box plot of economic cost",
        ylab="Property and Crop Damage")
```



The weather event that has the greatest economic consequence is:

```
print(health_res[which.max(econ_res$x),1])
```

```
## [1] "TORNADOES, TSTM WIND, HAIL"
```