The impact of weather events on population health and economic consequences in the U.S.

Synopsis

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

Questions:

1. Across the United States, which types of events (EVTYPE) are most harmful with respect to population health?
2. Across the United States, which types of events have the greatest economic consequences?

Data

The data for this assignment come in the form of a comma-separated-value file compressed via the bzip2 algorithm to reduce its size.

Dataset: Storm data

Documentation of how some of the variables are constructed/defined:

- National Weather Service (Storm Data Documentation, https://d396qusza40orc.cloudfront.net/repdata%2Fpeer2_doc%2Fpd01016005curr.pdf)
- National Climatic Data Center Storm Events (FAQ, https://d396qusza40orc.cloudfront.net/repdata%2Fpeer2_doc%2FNCDC%20Storm%20Events-FAQ%20Page.pdf)

The events in the database start in the year 1950 and end in November 2011. In the earlier years of the database there are generally fewer events recorded, most likely due to a lack of good records. More recent years should be considered more complete.

The full database consists of 902297 observations of 37 variables. Of these the principal data required to evaluate the economic and health consequences of various weather events are:

- EVTYPE - a factor variable giving the event type (e.g. tornado, flood, etc.)
- FATALITIES - a numerical variable of the number of fatalities
- INJURIES - a numerical variable of the number of injuries
- PROPDMG - a numerical variable giving the mantissa for the value of property damage in USD
- PROPDMGEXP - a factor variable giving the exponent for the value of property damage in USD
- CROPDMG - a numerical variable giving the mantissa for the value of crop damage in USD
- CROPDMGEXP - a factor variable giving the exponent for the value of crop damage in USD

Loading libraries

```
library(knitr)
library(ggplot2)
```

Load data

```
fileurl <- "https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2"
download.file(fileurl, destfile="stormdata", method="curl")
storm <- read.csv("stormdata", header=TRUE, sep=",")
```

Data processing

```
# Finding the total harm on population health with the sum of FATALITIES and
#INJURIES, and creating a new variable called, "health"
storm$health <- storm$FATALITIES+storm$INJURIES

# Calculating the no. of health harmed by EVTYPE
health <- aggregate(health ~ EVTYPE, data=storm, FUN=sum, na.rm=TRUE)

# Let's see the top 10 most severe weather events with the highest number of
# fatalities and injuries
top10health <- health[order(health$health, decreasing=TRUE),][1:10,]
print(top10health)
```

```
##                EVTYPE health
## 834           TORNADO  96979
## 130    EXCESSIVE HEAT   8428
## 856         TSTM WIND   7461
## 170             FLOOD   7259
## 464         LIGHTNING   6046
## 275              HEAT   3037
## 153       FLASH FLOOD   2755
## 427         ICE STORM   2064
## 760 THUNDERSTORM WIND   1621
## 972      WINTER STORM   1527
```

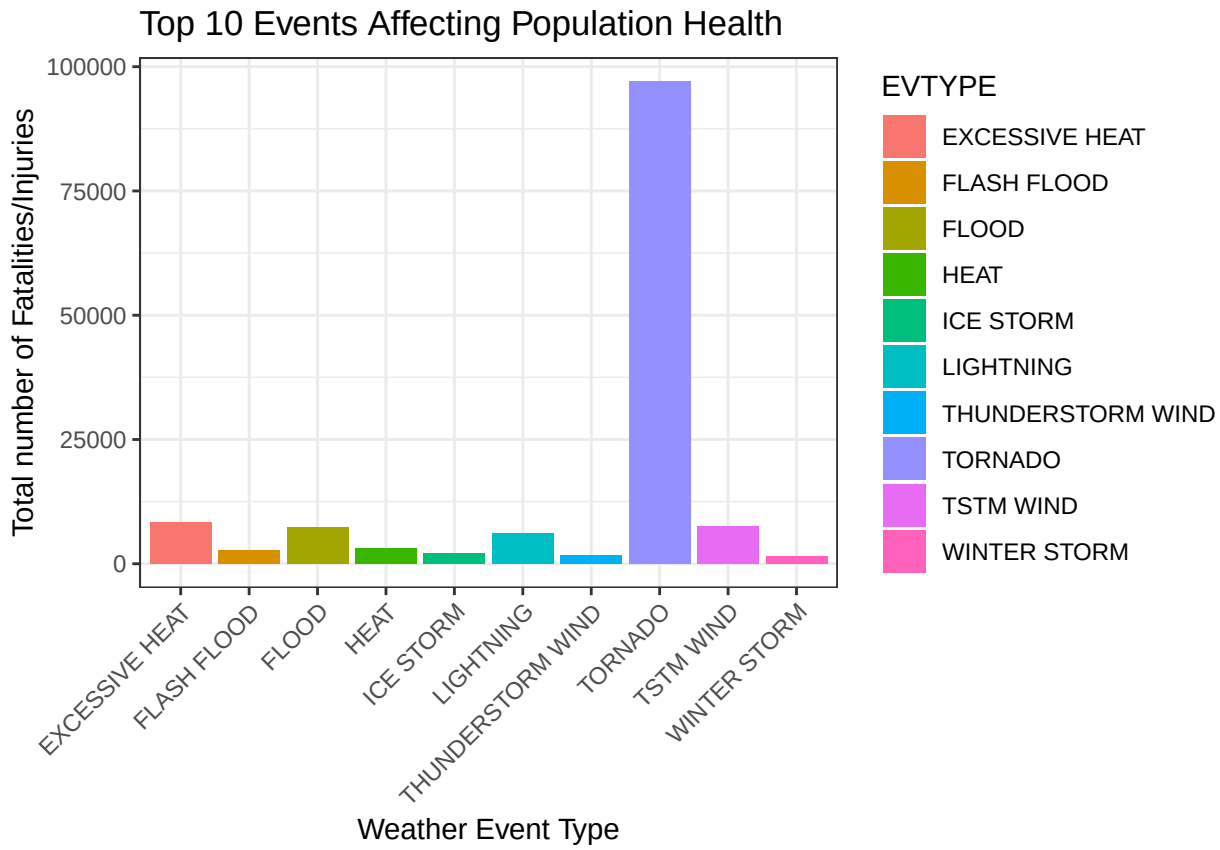
```
# Finding the total damage on ecomony with the sum of relevant CROP and PROP variables,
#and creating a new variable called, "econ"
EXP_cha <- c("B" , "M" , "K", "","m","0","1","2","3","4","5","6","7","8","+","-","H","h","?")
EXP_num <- c(10^9, 10^6, 10^3, 0,10^6,10,10,10,10,10,10,10,10,10,1,0,100,100,0)
storm$PROPDMG2 <- storm$PROPDMG * EXP_num[match(storm$PROPDMGEXP, EXP_cha)]
storm$CROPDMG2 <- storm$CROPDMG * EXP_num[match(storm$CROPDMGEXP, EXP_cha)]
storm$econ <- storm$PROPDMG2 + storm$CROPDMG2

# Calculating the economy harmed by EVTYPE
econ <- aggregate(econ ~ EVTYPE, data=storm, FUN=sum, na.rm=TRUE)

# Top 10 most severe weather events with the highest number of property
# and crop damages
top10dmg <- econ[order(econ$econ, decreasing=TRUE),][1:10,]
print(top10dmg)
```

```
##                EVTYPE         econ
## 170             FLOOD 150319678250
## 411 HURRICANE/TYPHOON  71913712800
## 834           TORNADO  57352117607
```

```
## 670       STORM SURGE  43323541000
## 244              HAIL  18757611527
## 153       FLASH FLOOD  17562132111
## 95            DROUGHT  15018672000
## 402         HURRICANE  14610229010
## 590       RIVER FLOOD  10148404500
## 427         ICE STORM   8967041810
```

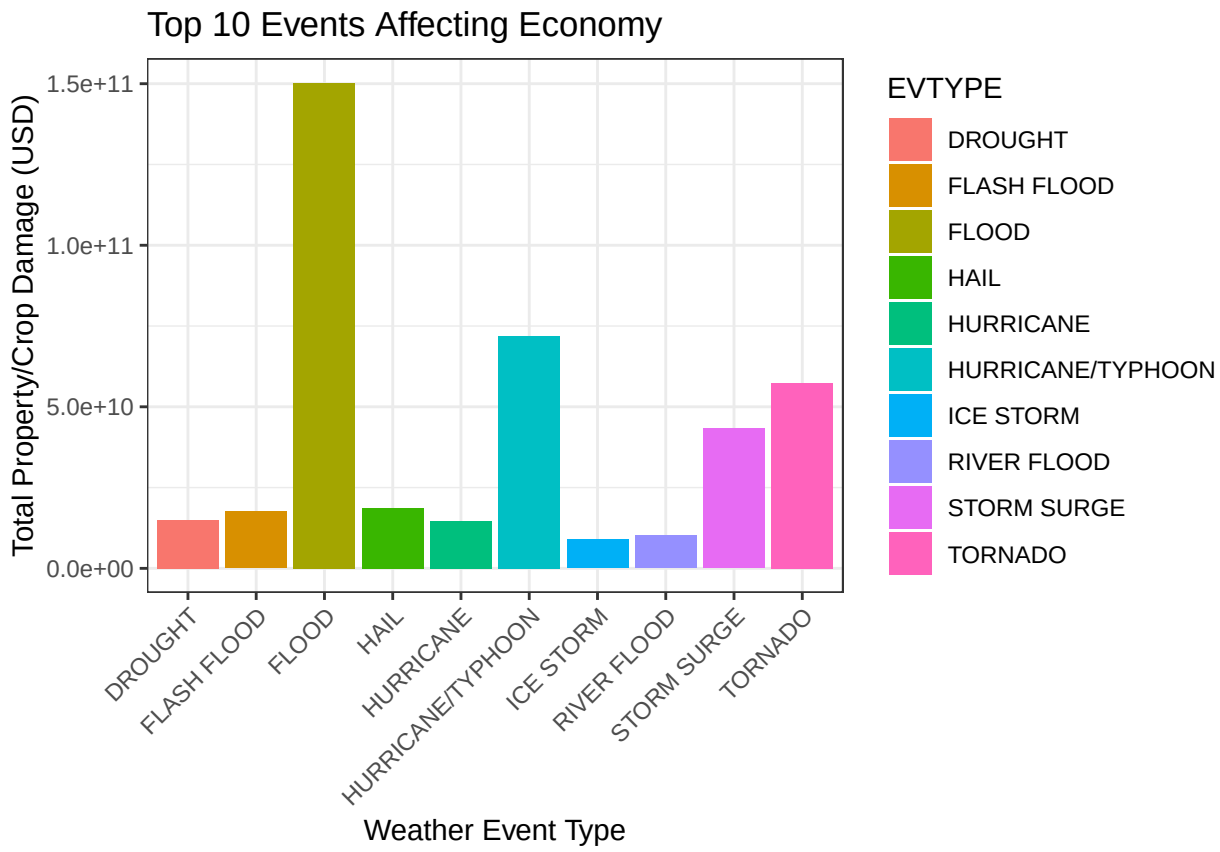
Results

```
# Population health (Fatalities + Injuries) harmed by severe weather events in the U.S.
ggplot(top10health)+aes(x=EVTYPE, y=health, fill=EVTYPE) +
    geom_bar(stat="identity") +
    labs(title = "Top 10 Events Affecting Population Health", x = "Weather Event Type",
    y = "Total number of Fatalities/Injuries") +
    theme_bw() +
    theme(axis.text.x=element_text(angle=45,hjust=1))
```




```
# Economic consequences harmed by property and crop damages in the U.S.
ggplot(top10dmg)+aes(x=EVTYPE, y=econ, fill=EVTYPE) +
    geom_bar(stat="identity") +
    labs(title = "Top 10 Events Affecting Economy", x = "Weather Event Type",
    y = "Total Property/Crop Damage (USD)") +
```

```
    theme_bw() +
    theme(axis.text.x=element_text(angle=45,hjust=1))
```



Conclusions

Across the United States, most fatalties and injuries were caused by Tornado; while the greatest economic damage (property and crop damages) was caused by flood.