

Reproducible Research Course Project/Storm Data

ludimila

05/11/2020

1 - Synopsis

The basic goal of this assignment is to explore the NOAA Storm Database and answer some basic questions about severe weather events. You must use the database to answer the questions below and show the code for your entire analysis.

Questions Your data analysis must address the following questions:

- Across the United States, which types of events (as indicated in the EVTYPE variable) are most harmful with respect to population health?
- Across the United States, which types of events have the greatest economic consequences?

2 - Data Processing

```
library(ggplot2)
library(plyr)
```

```
## Warning: package 'plyr' was built under R version 4.0.3
```

```
library(dplyr)
```

```
##
```

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

```
## The following objects are masked from 'package:plyr':
```

```
##
```

```
##      arrange, count, desc, failwith, id, mutate, rename, summarise,
##      summarize
```

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

```
##
```

```
##      filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      intersect, setdiff, setequal, union
```

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

```
## Warning: package 'R.utils' was built under R version 4.0.3

## Loading required package: R.oo

## Warning: package 'R.oo' was built under R version 4.0.3

## Loading required package: R.methodsS3

## Warning: package 'R.methodsS3' was built under R version 4.0.3

## R.methodsS3 v1.8.1 (2020-08-26 16:20:06 UTC) successfully loaded. See ?R.methodsS3 for help.

## R.oo v1.24.0 (2020-08-26 16:11:58 UTC) successfully loaded. See ?R.oo for help.

##
## Attaching package: 'R.oo'

## The following object is masked from 'package:R.methodsS3':
##
##      throw

## The following objects are masked from 'package:methods':
##
##      getClasses, getMethods

## The following objects are masked from 'package:base':
##
##      attach, detach, load, save

## R.utils v2.10.1 (2020-08-26 22:50:31 UTC) 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, nullfile, parse,
##      warnings

if(!file.exists("stormData.csv.bz2")) {
  download.file("https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2",
               destfile = "stormData.csv.bz2", method = "curl")
}

stormDATA <- read.csv(bzfile("stormData.csv.bz2"), sep = ",", header = T)
```

3 - Results

3.1 - Across the United States, which types of events are most harmful with respect to population health?

```
fatalitiesDATA <- aggregate(stormDATA$FATALITIES, by=list(EVTYPE=stormDATA$EVTYPE), FUN=sum)
fDATA <- head(fatalitiesDATA[order(-fatalitiesDATA$x),])

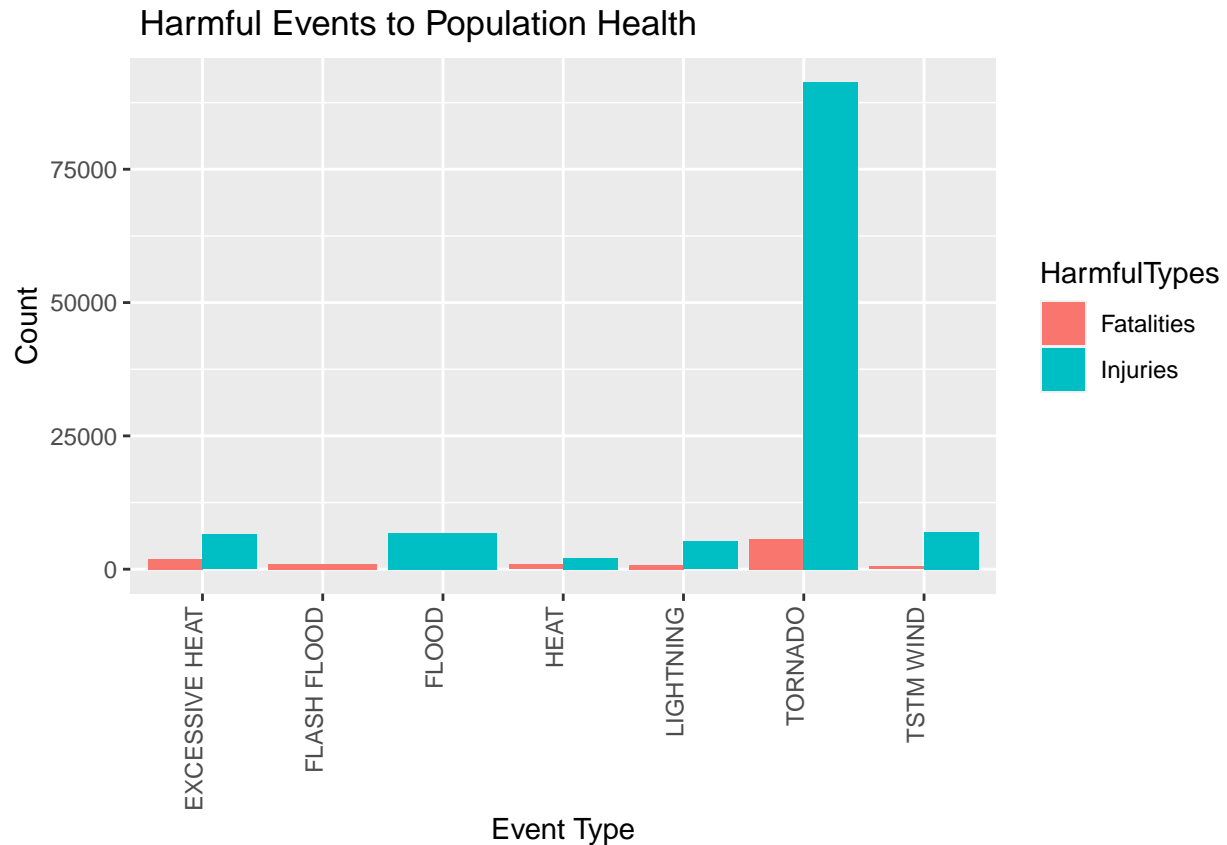
injuriesDATA <- aggregate(stormDATA$INJURIES, by=list(EVTYPE=stormDATA$EVTYPE), FUN=sum)
iDATA <- head(injuriesDATA[order(-injuriesDATA$x),])

# Merging Data
fDATA$HarmfulTypes <- "Fatalities"
iDATA$HarmfulTypes <- "Injuries"
mostHarm <- rbind(fDATA, iDATA)
print(mostHarm)
```

```
##           EVTYPE      x HarmfulTypes
## 834      TORNADO  5633  Fatalities
## 130 EXCESSIVE HEAT  1903  Fatalities
## 153    FLASH FLOOD   978  Fatalities
## 275         HEAT   937  Fatalities
## 464    LIGHTNING   816  Fatalities
## 856    TSTM WIND   504  Fatalities
## 8341     TORNADO 91346   Injuries
## 8561    TSTM WIND 6957   Injuries
## 170       FLOOD  6789   Injuries
## 1301 EXCESSIVE HEAT 6525   Injuries
## 4641    LIGHTNING  5230   Injuries
## 2751         HEAT  2100   Injuries
```

Plot below shows the results of the analysis

```
plot1 <- ggplot(mostHarm, aes(EVTYPE,x, fill = HarmfulTypes)) + geom_bar(position = "dodge", stat = "id")
  labs(title = " Harmful Events to Population Health", x = "Event Type", y = "Count") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0))
print(plot1)
```



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

```
unique(stormDATA$PROPDGMGEXP)
```

```
## [1] "K" "M" "" "B" "m" "+" "0" "5" "6" "?" "4" "2" "3" "h" "7" "H" "-" "1" "8"
```

```
unique(stormDATA$CROPDGMGEXP)
```

```
## [1] "" "M" "K" "m" "B" "?" "0" "k" "2"
```

```
PROPDGMG <- mapvalues(stormDATA$PROPDGMGEXP,
  c("K", "M", "", "B", "m", "+", "0", "5", "6", "?", "4", "2", "3", "h", "7", "H", "-", "1", "8"),
  c(1e3, 1e6, 1, 1e9, 1e6, 1, 1, 1e5, 1e6, 1, 1e4, 1e2, 1e3, 1, 1e7, 1e2, 1, 10, 1e8))

CROPDGMG <- mapvalues(stormDATA$CROPDGMGEXP,
  c("", "M", "K", "m", "B", "?", "0", "k", "2"),
  c(1, 1e6, 1e3, 1e6, 1e9, 1, 1, 1e3, 1e2))

stormDATA$PROPTOTALDMG <- as.numeric(PROPDGMG) * stormDATA$PROPDGMG
stormDATA$CROPTOTALDMG <- as.numeric(CROPDGMG) * stormDATA$CROPDGMG
```

```

stormDATA$TOTALDMG <- stormDATA$PROPTOTALDMG + stormDATA$CROPTOTALDMG

PROPDGM <- aggregate(stormDATA$PROPTOTALDMG, by=list(EVTYPE=stormDATA$EVTYPE), FUN=sum)
PROPDGM <- head(PROPDGM[order(-PROPDGM$x),])

CROPDGM <- aggregate(stormDATA$CROPTOTALDMG, by=list(EVTYPE=stormDATA$EVTYPE), FUN=sum)
CROPDGM <- head(CROPDGM[order(-CROPDGM$x),])

TOTALDMG <- aggregate(stormDATA$TOTALDMG, by=list(EVTYPE=stormDATA$EVTYPE), FUN=sum)
TOTALDMG <- head(TOTALDMG[order(-TOTALDMG$x),])

# Merging Data
PROPDGM$Type <- "Property Damage Caused"
CROPDGM$Type <- "Crop Damage Caused"
TOTALDMG$Type <- "Total Damage"
mostHarmECO <- rbind(PROPDGM, CROPDGM, TOTALDMG)
print(mostHarmECO)

```

##	EVTYPE	x	Type
## 170	FLOOD	144657709807	Property Damage Caused
## 411	HURRICANE/TYPHOON	69305840000	Property Damage Caused
## 834	TORNADO	56947380677	Property Damage Caused
## 670	STORM SURGE	43323536000	Property Damage Caused
## 153	FLASH FLOOD	16822673979	Property Damage Caused
## 244	HAIL	15735267513	Property Damage Caused
## 95	DROUGHT	13972566000	Crop Damage Caused
## 1701	FLOOD	5661968450	Crop Damage Caused
## 590	RIVER FLOOD	5029459000	Crop Damage Caused
## 427	ICE STORM	5022113500	Crop Damage Caused
## 2441	HAIL	3025954473	Crop Damage Caused
## 402	HURRICANE	2741910000	Crop Damage Caused
## 1702	FLOOD	150319678257	Total Damage
## 4111	HURRICANE/TYPHOON	71913712800	Total Damage
## 8341	TORNADO	57362333947	Total Damage
## 6701	STORM SURGE	43323541000	Total Damage
## 2442	HAIL	18761221986	Total Damage
## 1531	FLASH FLOOD	18243991079	Total Damage

Plot below shows the results of the analysis

```

plot2 <- ggplot(mostHarmECO, aes(EVTYPE,x), grpop=type) + geom_bar(aes(colour=Type, fill= Type),position
  labs(title = "Economically Harmful Events", x = "Event Type", y = "Count in $") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0))
print(plot2)

```

