Exploring NOAA

Masaaki Inaba

2015/01/25

Synopsis

This project involves exploring the U.S. National Oceanic and Atmospheric Administration's (NOAA) storm database and answer two questions.

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

Data Processing

```
options(scipen=100)

# Load data
setwd("~/Coursera/ReproducibleResearch/PeerAssessment2")
noaaData <- read.csv(bzfile("repdata_data_StormData.csv.bz2"))

# Extract columns related population health and economic consequences.
#  - EVTYPE:     Event type
#  - FATALITIES: Fatalities
#  - INJURIES:   Injuries
#  - PROPDMG:    The estimated amount of damage to property incurred by the weather event.
#  - CROPDMG:    The estimated amount of damage to crops incurred by the weather event.
noaaData <- noaaData[c("EVTYPE", "FATALITIES", "INJURIES", "PROPDMG", "CROPDMG")]

# Calculate mean values per event type.
noaaData <- aggregate(
  list(
    FATALITIES = noaaData$FATALITIES,
    INJURIES   = noaaData$INJURIES,
    PROPDMG    = noaaData$PROPDMG,
    CROPDMG    = noaaData$CROPDMG),
  by = list(EVTYPE = noaaData$EVTYPE),
  FUN = mean)

# Calculate harmfulness (assumed that it's defined as FATALITIES + INJURIES)
# and economic consequences (assumed that it's defined as PROPDMG + CROPDMG).
noaaData <- transform(noaaData, "HEALTH"   = noaaData$FATALITIES + noaaData$INJURIES)
noaaData <- transform(noaaData, "ECONOMIC" = noaaData$PROPDMG    + noaaData$CROPDMG)

# Extract top 10
health   <- head(noaaData[order(noaaData$HEALTH,   decreasing = TRUE),], n = 10)
economic <- head(noaaData[order(noaaData$ECONOMIC, decreasing = TRUE),], n = 10)
rownames(health)   <- health$EVTYPE
rownames(economic) <- economic$EVTYPE
health   <- health[c("FATALITIES", "INJURIES", "HEALTH")]
economic <- economic[c("PROPDMG", "CROPDMG", "ECONOMIC")]
rm(noaaData)
```

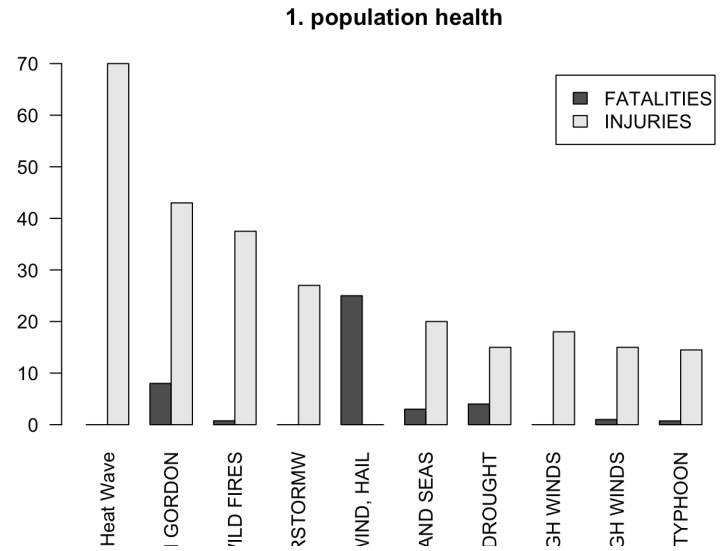
Results

1. Most harmful type of event is below with respect to population health when harmfulness is defined as (number of fatalities + number of injuries).

```
rownames(health[1,])
```

```
## [1] "Heat Wave"
```

```
barplot(
  t(data.matrix(health[c("FATALITIES", "INJURIES")])),
  beside = TRUE,
  main = "1. population health",
  legend = c("FATALITIES", "INJURIES"),
  las = 2)
```



2. The type of events have the greatest economic consequences is below.

```
rownames(economic[1,])
```

```
## [1] "TROPICAL STORM GORDON"
```

```
barplot(
  t(data.matrix(economic[c("PROPDMG", "CROPDMG")])),
  beside = TRUE,
  main = "2. economic consequences",
  legend = c("PROPDMG", "CROPDMG"),
  las = 2,
  font = 1)
```